

DETECTION OF RNA

The present application is a continuation-in-part of U.S. Patent Application No. 09/577,304 and 09/758,282 and is a continuation-in-part of U.S. Patent Application No. 09/350,309, which is a divisional application U.S. Patent No. 5,985,557, and is also a continuation-in-part of co-pending application serial no. 09/381,212 which is a national entry of PCT application no. US 98/05809, which claims priority to U.S. Patent Nos. 5,994,069, 6,090,543, 5,985,557, 6,001,567, and 5,846,717 and PCT application no. US 97/01072.

FIELD OF THE INVENTION

The present invention relates to compositions and methods for the detection and characterization of nucleic acid sequences and variations in nucleic acid sequences. The present invention relates to methods for forming a nucleic acid cleavage structure on a target sequence and cleaving the nucleic acid cleavage structure in a site-specific manner. For example, in some embodiments, the 5' nuclease activity of a variety of enzymes is used to cleave the target-dependent cleavage structure, thereby indicating the presence of specific nucleic acid sequences or specific variations thereof.

BACKGROUND OF THE INVENTION

Methods for the detection and characterization of specific nucleic acid sequences and sequence variations have been used to detect the presence of viral or bacterial nucleic acid sequences indicative of an infection, to detect the presence of variants or alleles of genes associated with disease and cancers. These methods also find application in the identification of sources of nucleic acids, as for forensic analysis or for paternity determinations.

Various methods are known to the art that may be used to detect and characterize specific nucleic acid sequences and sequence variants. Nonetheless, with the completion of the nucleic acid sequencing of the human genome, as well as the genomes of numerous pathogenic organisms, the demand for fast, reliable, cost-effective and user-friendly tests for the detection of specific nucleic acid sequences continues to grow. Importantly, these

tests must be able to create a detectable signal from samples that contain very few copies of the sequence of interest. The following discussion examines two levels of nucleic acid detection assays currently in use: I. Signal Amplification Technology for detection of rare sequences II. Direct Detection Technology for quantitative detection of sequences, and III. Direct Detection of RNA.

I. Signal Amplification Technology Methods For Amplification

The "Polymerase Chain Reaction" (PCR) comprises the first generation of methods for nucleic acid amplification. However, several other methods have been developed that employ the same basis of specificity, but create signal by different amplification mechanisms. These methods include the "Ligase Chain Reaction" (LCR), "Self-Sustained Synthetic Reaction" (3SR/NASBA), and "Q β -Replicase" (Q β).

Polymerase Chain Reaction (PCR)

The polymerase chain reaction (PCR), as described in U.S. Patent Nos. 4,683,195, 4,683,202, and 4,965,188 to Mullis and Mullis *et al.* (the disclosures of which are hereby incorporated by reference), describe a method for increasing the concentration of a segment of target sequence in a mixture of genomic DNA without cloning or purification. This technology provides one approach to the problems of low target sequence concentration. PCR can be used to directly increase the concentration of the target to an easily detectable level. This process for amplifying the target sequence involves introducing a molar excess of two oligonucleotide primers that are complementary to their respective strands of the double-stranded target sequence to the DNA mixture containing the desired target sequence. The mixture is denatured and then allowed to hybridize. Following hybridization, the primers are extended with polymerase so as to form complementary strands. The steps of denaturation, hybridization, and polymerase extension can be repeated as often as needed, in order to obtain relatively high concentrations of a segment of the desired target sequence.

The length of the segment of the desired target sequence is determined by the relative positions of the primers with respect to each other, and, therefore, this length is a

controllable parameter. Because the desired segments of the target sequence become the dominant sequences (in terms of concentration) in the mixture, they are said to be "PCR-amplified."

Ligase Chain Reaction (LCR or LAR)

5 The ligase chain reaction (LCR; sometimes referred to as "Ligase Amplification Reaction" (LAR) described by Barany, Proc. Natl. Acad. Sci., 88:189 (1991); Barany, PCR Methods and Applic., 1:5 (1991); and Wu and Wallace, Genomics 4:560 (1989) has developed into a well-recognized alternative method for amplifying nucleic acids. In LCR, four oligonucleotides, two adjacent oligonucleotides that uniquely hybridize to one
10 strand of target DNA, and a complementary set of adjacent oligonucleotides, that hybridize to the opposite strand are mixed and DNA ligase is added to the mixture. Provided that there is complete complementarity at the junction, ligase will covalently link each set of hybridized molecules. Importantly, in LCR, two probes are ligated together only when they base-pair with sequences in the target sample, without gaps or
15 mismatches. Repeated cycles of denaturation, hybridization and ligation amplify a short segment of DNA. LCR has also been used in combination with PCR to achieve enhanced detection of single-base changes. Segev, PCT Public. No. W09001069 A1 (1990). However, because the four oligonucleotides used in this assay can pair to form two short ligatable fragments, there is the potential for the generation of target-independent
20 background signal. The use of LCR for mutant screening is limited to the examination of specific nucleic acid positions.

Self-Sustained Synthetic Reaction (3SR/NASBA)

 The self-sustained sequence replication reaction (3SR) (Guatelli *et al.*, Proc. Natl. Acad. Sci., 87:1874-1878 [1990], with an erratum at Proc. Natl. Acad. Sci., 87:7797
25 [1990]) is a transcription-based *in vitro* amplification system (Kwok *et al.*, Proc. Natl. Acad. Sci., 86:1173-1177 [1989]) that can exponentially amplify RNA sequences at a uniform temperature. The amplified RNA can then be utilized for mutation detection (Fahy *et al.*, PCR Meth. Appl., 1:25 [1991]). In this method, an oligonucleotide primer is used to add a phage RNA polymerase promoter to the 5' end of the sequence of interest.
30 In a cocktail of enzymes and substrates that includes a second primer, reverse

transcriptase, RNase H, RNA polymerase and ribo-and deoxyribonucleoside triphosphates, the target sequence undergoes repeated rounds of transcription, cDNA synthesis and second-strand synthesis to amplify the area of interest. The use of 3SR to detect mutations is kinetically limited to screening small segments of DNA (*e.g.*, 200-300 base pairs).

Q-Beta (Q β) Replicase

In this method, a probe that recognizes the sequence of interest is attached to the replicatable RNA template for Q β replicase. A previously identified major problem with false positives resulting from the replication of unhybridized probes has been addressed through use of a sequence-specific ligation step. However, available thermostable DNA ligases are not effective on this RNA substrate, so the ligation must be performed by T4 DNA ligase at low temperatures (37°C). This prevents the use of high temperature as a means of achieving specificity as in the LCR, the ligation event can be used to detect a mutation at the junction site, but not elsewhere.

Table 2 below, lists some of the features desirable for systems useful in sensitive nucleic acid diagnostics, and summarizes the abilities of each of the major amplification methods (*See also*, Landgren, Trends in Genetics 9:199 [1993]).

A successful diagnostic method must be very specific. A straight-forward method of controlling the specificity of nucleic acid hybridization is by controlling the temperature of the reaction. While the 3SR/NASBA, and Q β systems are all able to generate a large quantity of signal, one or more of the enzymes involved in each cannot be used at high temperature (*i.e.*, >55°C). Therefore the reaction temperatures cannot be raised to prevent non-specific hybridization of the probes. If probes are shortened in order to make them melt more easily at low temperatures, the likelihood of having more than one perfect match in a complex genome increases. For these reasons, PCR and LCR currently dominate the research field in detection technologies.

TABLE 1

| Feature | Method | | | | |
|--|--------|-----|--------------|--------------|-----------|
| | PCR | LCR | PCR & LCR | 3SR NASBA | Q β |
| Amplifies Target | + | + | + | + | |
| Recognition of Independent Sequences Required | + | + | + | + | + |
| Performed at High Temp. | + | + | | | |
| Operates at Fixed Temp. | | | | + | + |
| Exponential Amplification | + | + | + | + | + |
| Generic Signal Generation | | | | | + |
| Easily Automatable | | | | | |

The basis of the amplification procedure in the PCR and LCR is the fact that the products of one cycle become usable templates in all subsequent cycles, consequently doubling the population with each cycle. The final yield of any such doubling system can be expressed as: $(1+X)^n = y$, where "X" is the mean efficiency (percent copied in each cycle), "n" is the number of cycles, and "y" is the overall efficiency, or yield of the reaction (Mullis, PCR Methods Applic., 1:1 [1991]). If every copy of a target DNA is utilized as a template in every cycle of a polymerase chain reaction, then the mean efficiency is 100%. If 20 cycles of PCR are performed, then the yield will be 2^{20} , or 1,048,576 copies of the starting material. If the reaction conditions reduce the mean efficiency to 85%, then the yield in those 20 cycles will be only 1.85^{20} , or 220,513 copies of the starting material. In other words, a PCR running at 85% efficiency will yield only 21% as much final product, compared to a reaction running at 100% efficiency. A

reaction that is reduced to 50% mean efficiency will yield less than 1% of the possible product.

In practice, routine polymerase chain reactions rarely achieve the theoretical maximum yield, and PCRs are usually run for more than 20 cycles to compensate for the lower yield. At 50% mean efficiency, it would take 34 cycles to achieve the million-fold amplification theoretically possible in 20, and at lower efficiencies, the number of cycles required becomes prohibitive. In addition, any background products that amplify with a better mean efficiency than the intended target will become the dominant products.

Also, many variables can influence the mean efficiency of PCR, including target DNA length and secondary structure, primer length and design, primer and dNTP concentrations, and buffer composition, to name but a few. Contamination of the reaction with exogenous DNA (*e.g.*, DNA spilled onto lab surfaces) or cross-contamination is also a major consideration. Reaction conditions must be carefully optimized for each different primer pair and target sequence, and the process can take days, even for an experienced investigator. The laboriousness of this process, including numerous technical considerations and other factors, presents a significant drawback to using PCR in the clinical setting. Indeed, PCR has yet to penetrate the clinical market in a significant way. The same concerns arise with LCR, as LCR must also be optimized to use different oligonucleotide sequences for each target sequence. In addition, both methods require expensive equipment, capable of precise temperature cycling.

Many applications of nucleic acid detection technologies, such as in studies of allelic variation, involve not only detection of a specific sequence in a complex background, but also the discrimination between sequences with few, or single, nucleotide differences. One method for the detection of allele-specific variants by PCR is based upon the fact that it is difficult for *Taq* polymerase to synthesize a DNA strand when there is a mismatch between the template strand and the 3' end of the primer. An allele-specific variant may be detected by the use of a primer that is perfectly matched with only one of the possible alleles; the mismatch to the other allele acts to prevent the extension of the primer, thereby preventing the amplification of that sequence. This method has a substantial limitation in that the base composition of the mismatch influences the ability to prevent extension across the mismatch, and certain mismatches

do not prevent extension or have only a minimal effect (Kwok *et al.*, Nucl. Acids Res., 18:999 [1990]).)

A similar 3'-mismatch strategy is used with greater effect to prevent ligation in the LCR (Barany, PCR Meth. Applic., 1:5 [1991]). Any mismatch effectively blocks the action of the thermostable ligase, but LCR still has the drawback of target-independent background ligation products initiating the amplification. Moreover, the combination of PCR with subsequent LCR to identify the nucleotides at individual positions is also a clearly cumbersome proposition for the clinical laboratory.

II. Direct Detection Technology

When a sufficient amount of a nucleic acid to be detected is available, there are advantages to detecting that sequence directly, instead of making more copies of that target, (*e.g.*, as in PCR and LCR). Most notably, a method that does not amplify the signal exponentially is more amenable to quantitative analysis. Even if the signal is enhanced by attaching multiple dyes to a single oligonucleotide, the correlation between the final signal intensity and amount of target is direct. Such a system has an additional advantage that the products of the reaction will not themselves promote further reaction, so contamination of lab surfaces by the products is not as much of a concern. Traditional methods of direct detection including Northern and Southern blotting and RNase protection assays usually require the use of radioactivity and are not amenable to automation. Recently devised techniques have sought to eliminate the use of radioactivity and/or improve the sensitivity in automatable formats. Two examples are the "Cycling Probe Reaction" (CPR), and "Branched DNA" (bDNA)

The cycling probe reaction (CPR) (Duck *et al.*, BioTech., 9:142 [1990]), uses a long chimeric oligonucleotide in which a central portion is made of RNA while the two termini are made of DNA. Hybridization of the probe to a target DNA and exposure to a thermostable RNase H causes the RNA portion to be digested. This destabilizes the remaining DNA portions of the duplex, releasing the remainder of the probe from the target DNA and allowing another probe molecule to repeat the process. The signal, in the form of cleaved probe molecules, accumulates at a linear rate. While the repeating

process increases the signal, the RNA portion of the oligonucleotide is vulnerable to RNases that may be carried through sample preparation.

Branched DNA (bDNA), described by Urdea *et al.*, Gene 61:253-264 (1987), involves oligonucleotides with branched structures that allow each individual
5 oligonucleotide to carry 35 to 40 labels (*e.g.*, alkaline phosphatase enzymes). While this enhances the signal from a hybridization event, signal from non-specific binding is similarly increased.

While both of these methods have the advantages of direct detection discussed above, neither the CPR or bDNA methods can make use of the specificity allowed by the
10 requirement of independent recognition by two or more probe (oligonucleotide) sequences, as is common in the signal amplification methods described in Section I. above. The requirement that two oligonucleotides must hybridize to a target nucleic acid in order for a detectable signal to be generated confers an extra measure of stringency on any detection assay. Requiring two oligonucleotides to bind to a target nucleic acid
15 reduces the chance that false "positive" results will be produced due to the non-specific binding of a probe to the target. The further requirement that the two oligonucleotides must bind in a specific orientation relative to the target, as is required in PCR, where oligonucleotides must be oppositely but appropriately oriented such that the DNA polymerase can bridge the gap between the two oligonucleotides in both directions,
20 further enhances specificity of the detection reaction. However, it is well known to those in the art that even though PCR utilizes two oligonucleotide probes (termed primers) "non-specific" amplification (*i.e.*, amplification of sequences not directed by the two primers used) is a common artifact. This is in part because the DNA polymerase used in PCR can accommodate very large distances, measured in nucleotides, between the
25 oligonucleotides and thus there is a large window in which non-specific binding of an oligonucleotide can lead to exponential amplification of inappropriate product. The LCR, in contrast, cannot proceed unless the oligonucleotides used are bound to the target adjacent to each other and so the full benefit of the dual oligonucleotide hybridization is realized.

An ideal direct detection method would combine the advantages of the direct detection assays (*e.g.*, easy quantification and minimal risk of carry-over contamination) with the specificity provided by a dual oligonucleotide hybridization assay.

5 **III. Direct Detection of RNA**

In molecular medicine, a simple and cost-effective method for direct and quantitative RNA detection would greatly facilitate the analysis of RNA viruses and the measurement of specific gene expression. Both of these issues are currently pressing problems in the field. Despite this need, few techniques have emerged that are truly
10 direct. PCR-based detection assays require conversion of RNA to DNA by reverse transcriptase before amplification, introducing a variable that can compromise accurate quantification. Furthermore, PCR and other methods based on exponential amplification (*e.g.*, NASBA) require painstaking containment measures to avoid cross-contamination, and have difficulty distinguishing small differences (*e.g.*, 2 to 3-fold) in quantity. Other
15 tests that directly examine RNA suffer from a variety of drawbacks, including time consuming autoradiography steps (*e.g.*, RNase protection assays), or overnight reaction times (*e.g.*, branched DNA assays). With over 1.5 million viral load measurements being performed in the U.S. every year, there is clearly an enormous potential for an inexpensive, rapid, high-throughput system for the quantitative measurement of RNA.

20 Techniques for direct, quantitative detection of mRNA are vital for monitoring expression of a number of different genes. In particular, levels of cytokine expression (*e.g.*, interleukins and lymphokines) are being exploited as clinical measures of immune response in the progression of a wide variety of diseases (Van Deuren *et al.*, J. Int. Fed. Clin. Chem., 5:216 [1993], Van Deuren *et al.*, J. Inf. Dis., 169:157 [1994], Perenboom *et al.*, Eur. J. Clin. Invest., 26:159 [1996], Guidotti *et al.*, Immunity 4:25 [1996]) as well as
25 in monitoring transplant recipients (Grant *et al.*, Transplantation 62:910 [1996]).

Additionally, the monitoring of viral load and identification of viral genotype have great clinical significance for individuals suffering viral infections by such pathogens as HIV or Hepatitis C virus (HCV). There is a high correlation between viral load (*i.e.*, the
30 absolute number of viral particles in the bloodstream) and time to progression to AIDS (Mellors *et al.*, Science 272:1167 [1996], Saag *et al.*, Nature Medicine 2:625 [1996]).

For that reason, viral load, as measured by quantitative nucleic acid based testing, is becoming a standard monitoring procedure for evaluating the efficacy of treatment and the clinical status of HIV positive patients. It is thought to be essential to reduce viral load as early in the course of infection as possible and to evaluate viral levels on a regular basis. In the case of HCV, viral genotype has great clinical significance, with correlations to severity of liver disease and responsiveness to interferon therapy. Furthermore, because HCV cannot be grown in culture, it is only by establishing correlations between characteristics like viral genotype and clinical outcome that new antiviral treatments can be evaluated.

While the above mentioned methods have been serviceable for low throughput, research applications, or for limited clinical application, it is clear that large scale quantitative analysis of RNA readily adaptable to any genetic system will require a more innovative approach. An ideal direct detection method would combine the advantages of the direct detection assays (*e.g.*, easy quantification and minimal risk of carry-over contamination) with the specificity provided by a dual oligonucleotide hybridization assay.

Many of the methods described above rely on hybridization alone to distinguish a target molecule from other nucleic acids. Although some of these methods can be highly sensitive, they often cannot quantitate and distinguish closely related mRNAs accurately, especially such RNAs expressed at different levels in the same sample. While the above-mentioned methods are serviceable for some purposes, a need exists for a technology that is particularly adept at distinguishing particular RNAs from closely related molecules.

SUMMARY OF THE INVENTION

The present invention relates to compositions and methods for the detection and characterization of nucleic acid sequences and variations in nucleic acid sequences. The present invention relates to methods for forming a nucleic acid cleavage structure on a target sequence and cleaving the nucleic acid cleavage structure in a site-specific manner.

For example, in some embodiments, the 5' nuclease activity of a variety of enzymes is

used to cleave the target-dependent cleavage structure, thereby indicating the presence of specific nucleic acid sequences or specific variations thereof.

The present invention provides structure-specific cleavage agents (e.g., nucleases) from a variety of sources, including mesophilic, psychrophilic, thermophilic, and hyperthermophilic organisms. The preferred structure-specific nucleases are thermostable. Thermostable structure-specific nucleases are contemplated as particularly useful in that they operate at temperatures where nucleic acid hybridization is extremely specific, allowing for allele-specific detection (including single-base mismatches). In one embodiment, the thermostable structure-specific nucleases are thermostable 5' nucleases comprising altered polymerases derived from the native polymerases of *Thermus* species, including, but not limited to *Thermus aquaticus*, *Thermus flavus*, and *Thermus thermophilus*. However, the invention is not limited to the use of thermostable 5' nucleases. Thermostable structure-specific nucleases from the FEN-1, RAD2 and XPG class of nucleases are also preferred.

The present invention provides a method for detecting a target sequence (e.g., a mutation, polymorphism, etc), comprising providing a sample suspected of containing the target sequence; oligonucleotides capable of forming an invasive cleavage structure in the presence of the target sequence; and an agent for detecting the presence of an invasive cleavage structure; and exposing the sample to the oligonucleotides and the agent. In some embodiments, the method further comprises the step of detecting a complex comprising the agent and the invasive cleavage structure (directly or indirectly). In some embodiments, the agent comprises a cleavage agent. In some preferred embodiments, the exposing of the sample to the oligonucleotides and the agent comprises exposing the sample to the oligonucleotides and the agent under conditions wherein an invasive cleavage structure is formed between the target sequence and the oligonucleotides if the target sequence is present in the sample, wherein the invasive cleavage structure is cleaved by the cleavage agent to form a cleavage product. In some embodiments, the method further comprises the step of detecting the cleavage product. In some embodiments, the target sequence comprises a first region and a second region, the second region downstream of and contiguous to the first region, and wherein the oligonucleotides comprise first and second oligonucleotides, wherein at least a portion of

the first oligonucleotide is completely complementary to the first portion of the target sequence and wherein the second oligonucleotide comprises a 3' portion and a 5' portion, wherein the 5' portion is completely complementary to the second portion of said target nucleic acid.

5 The present invention also provides a kit for detecting such target sequences, said kit comprising oligonucleotides capable of forming an invasive cleavage structure in the presence of the target sequence. In some embodiments, the kit further comprises an agent for detecting the presence of an invasive cleavage structure (e.g., a cleavage agent). In some embodiments, the oligonucleotides comprise first and second oligonucleotides, said
10 first oligonucleotide comprising a 5' portion complementary to a first region of the target nucleic acid and said second oligonucleotide comprising a 3' portion and a 5' portion, said 5' portion complementary to a second region of the target nucleic acid downstream of and contiguous to the first portion. In some preferred embodiments, the target sequence comprises

15 The present invention also provides methods for detecting the presence of a target nucleic acid molecule by detecting non-target cleavage products comprising providing: a cleavage agent; a source of target nucleic acid, the target nucleic acid comprising a first region and a second region, the second region downstream of and contiguous to the first region; a first oligonucleotide, wherein at least a portion of the first oligonucleotide is
20 completely complementary to the first portion of the target nucleic acid; and a second oligonucleotide comprising a 3' portion and a 5' portion, wherein the 5' portion is completely complementary to the second portion of the target nucleic acid; mixing the cleavage agent, the target nucleic acid, the first oligonucleotide and the second oligonucleotide to create a reaction mixture under reaction conditions such that at least
25 the portion of the first oligonucleotide is annealed to the first region of said target nucleic acid and wherein at least the 5' portion of the second oligonucleotide is annealed to the second region of the target nucleic acid so as to create a cleavage structure, and wherein cleavage of the cleavage structure occurs to generate non-target cleavage product; and detecting the cleavage of the cleavage structure.

30 The detection of the cleavage of the cleavage structure can be carried out in any manner. In some embodiments, the detection of the cleavage of the cleavage structure

comprises detecting the non-target cleavage product. In yet other embodiments, the detection of the cleavage of the cleavage structure comprises detection of fluorescence, mass, or fluorescence energy transfer. Other detection methods include, but are not limited to detection of radioactivity, luminescence, phosphorescence, fluorescence polarization, and charge. In some embodiments, detection is carried out by a method comprising providing the non-target cleavage product; a composition comprising two single-stranded nucleic acids annealed so as to define a single-stranded portion of a protein binding region; and a protein; and exposing the non-target cleavage product to the single-stranded portion of the protein binding region under conditions such that the protein binds to the protein binding region. In some embodiments, the protein comprises a nucleic acid producing protein, wherein the nucleic acid producing protein binds to the protein-binding region and produces nucleic acid. In some embodiments, the protein-binding region is a template-dependent RNA polymerase binding region (e.g., a T7 RNA polymerase binding region). In other embodiments, the detection is carried out by a method comprising providing the non-target cleavage product; a single continuous strand of nucleic acid comprising a sequence defining a single strand of an RNA polymerase binding region; a template-dependent DNA polymerase; and a template-dependent RNA polymerase; exposing the non-target cleavage product to the RNA polymerase binding region under conditions such that the non-target cleavage product binds to a portion of the single strand of the RNA polymerase binding region to produce a bound non-target cleavage product; exposing the bound non-target cleavage product to the template-dependent DNA polymerase under conditions such that a double-stranded RNA polymerase binding region is produced; and exposing the double-stranded RNA polymerase binding region to the template-dependent RNA polymerase under conditions such that RNA transcripts are produced. In some embodiments, the method further comprises the step of detecting the RNA transcripts. In some embodiments, the template-dependent RNA polymerase is T7 RNA polymerase.

The present invention is not limited by the nature of the 3' portion of the second oligonucleotide. In some preferred embodiments, the 3' portion of the second oligonucleotide comprises a 3' terminal nucleotide not complementary to the target

nucleic acid. In some embodiments, the 3' portion of the second oligonucleotide consists of a single nucleotide not complementary to the target nucleic acid.

Any of the components of the method may be attached to a solid support. For example, in some embodiments, the first oligonucleotide is attached to a solid support. In other embodiments, the second oligonucleotide is attached to a solid support.

The cleavage agent can be any agent that is capable of cleaving invasive cleavage structures. In some preferred embodiments, the cleavage agent comprises a structure-specific nuclease. In particularly preferred embodiments, the structure-specific nuclease comprises a thermostable structure-specific nuclease (e.g., a thermostable 5' nuclease).

Thermostable structure-specific nucleases include, but are not limited to, those having an amino acid sequence homologous to a portion of the amino acid sequence of a thermostable DNA polymerase derived from a thermophilic organism (e.g., *Thermus aquaticus*, *Thermus flavus*, and *Thermus thermophilus*). In other embodiments, the thermostable structure-specific nuclease comprises a nuclease from the FEN-1, RAD2 or XPG classes of nucleases, or chimerical structures containing one or more portions of any of the above cleavage agents.

The method is not limited by the nature of the target nucleic acid. In some embodiments, the target nucleic acid is single stranded or double stranded DNA or RNA. In some embodiments, double stranded nucleic acid is rendered single stranded (e.g., by heat) prior to formation of the cleavage structure. In some embodiment, the source of target nucleic acid comprises a sample containing genomic DNA. Sample include, but are not limited to, blood, saliva, cerebral spinal fluid, pleural fluid, milk, lymph, sputum and semen.

In some embodiments, the reaction conditions for the method comprise providing a source of divalent cations. In some preferred embodiments, the divalent cation is selected from the group comprising Mn^{2+} and Mg^{2+} ions. In some embodiments, the reaction conditions for the method comprise providing the first and the second oligonucleotides in concentration excess compared to the target nucleic acid.

In some embodiments, the method further comprises providing a third oligonucleotide complementary to a third portion of said target nucleic acid upstream of

the first portion of the target nucleic acid, wherein the third oligonucleotide is mixed with the reaction mixture.

The present invention also provides a method for detecting the presence of a target nucleic acid molecule by detecting non-target cleavage products comprising

5 providing: a cleavage agent; a source of target nucleic acid, the target nucleic acid comprising a first region and a second region, the second region downstream of and contiguous to the first region; a plurality of first oligonucleotides, wherein at least a portion of the first oligonucleotides is completely complementary to the first portion of the target nucleic acid; a second oligonucleotide comprising a 3' portion and a 5' portion,

10 wherein said 5' portion is completely complementary to the second portion of the target nucleic acid; mixing the cleavage agent, the target nucleic acid, the plurality of first oligonucleotides and second oligonucleotide to create a reaction mixture under reaction conditions such that at least the portion of a first oligonucleotide is annealed to the first region of the target nucleic acid and wherein at least the 5' portion of the second

15 oligonucleotide is annealed to the second region of the target nucleic acid so as to create a cleavage structure, and wherein cleavage of the cleavage structure occurs to generate non-target cleavage product, wherein the conditions permit multiple cleavage structures to form and be cleaved from the target nucleic acid; and detecting the cleavage of said cleavage structures. In some embodiments, the conditions comprise isothermal

20 conditions that permit the plurality of first oligonucleotides to dissociate from the target nucleic acid. While the present invention is limited by the number of cleavage structure formed on a particular target nucleic acid, in some preferred embodiments, two or more (3, 4, 5, . . . , 10, . . . , 10000, . . .) of the plurality of first oligonucleotides form cleavage structures with a particular target nucleic acid, wherein the cleavage structures are

25 cleaved to produce the non-target cleavage products.

The present invention also provides methods wherein a cleavage product from the above methods is used in a further invasive cleavage reaction. For example, the present invention provides a method comprising providing a cleavage agent; a first target nucleic acid, the first target nucleic acid comprising a first region and a second region, the second

30 region downstream of and contiguous to the first region; a first oligonucleotide, wherein at least a portion of the first oligonucleotide is completely complementary to the first

portion of the first target nucleic acid; a second oligonucleotide comprising a 3' portion and a 5' portion, wherein the 5' portion is completely complementary to the second portion of the first target nucleic acid; a second target nucleic acid, said second target nucleic acid comprising a first region and a second region, the second region downstream of and contiguous to the first region; and a third oligonucleotide, wherein at least a portion of the third oligonucleotide is completely complementary to the first portion of the second target nucleic acid; generating a first cleavage structure wherein at least said portion of the first oligonucleotide is annealed to the first region of the first target nucleic acid and wherein at least the 5' portion of the second oligonucleotide is annealed to the second region of the first target nucleic acid and wherein cleavage of the first cleavage structure occurs via the cleavage agent thereby cleaving the first oligonucleotide to generate a fourth oligonucleotide, said fourth oligonucleotide comprising a 3' portion and a 5' portion, wherein the 5' portion is completely complementary to the second portion of the second target nucleic acid; generating a second cleavage structure under conditions wherein at least said portion of the third oligonucleotide is annealed to the first region of the second target nucleic acid and wherein at least the 5' portion of the fourth oligonucleotide is annealed to the second region of the second target nucleic acid and wherein cleavage of the second cleavage structure occurs to generate a cleavage fragment; and detecting the cleavage of the second cleavage structure. In some preferred embodiments, the 3' portion of the fourth oligonucleotide comprises a 3' terminal nucleotide not complementary to the second target nucleic acid. In some embodiments, the 3' portion of the third oligonucleotide is covalently linked to the second target nucleic acid. In some embodiments, the second target nucleic acid further comprises a 5' region, wherein the 5' region of the second target nucleic acid is the third oligonucleotide.

The present invention further provides kits comprising: a cleavage agent; a first oligonucleotide comprising a 5' portion complementary to a first region of a target nucleic acid; and a second oligonucleotide comprising a 3' portion and a 5' portion, said 5' portion complementary to a second region of the target nucleic acid downstream of and contiguous to the first portion. In some embodiments, the 3' portion of the second oligonucleotide comprises a 3' terminal nucleotide not complementary to the target nucleic acid. In preferred embodiments, the 3' portion of the second oligonucleotide

consists of a single nucleotide not complementary to the target nucleic acid. In some embodiments, the kit further comprises a solid support. For example, in some embodiments, the first and/or second oligonucleotide is attached to said solid support. In some embodiments, the kit further comprises a buffer solution. In some preferred

5 embodiments, the buffer solution comprises a source of divalent cations (e.g., Mn^{2+} and/or Mg^{2+} ions). In some specific embodiments, the kit further comprises a third oligonucleotide complementary to a third portion of the target nucleic acid upstream of the first portion of the first target nucleic acid. In yet other embodiments, the kit further comprises a target nucleic acid. In some embodiments, the kit further comprises a second

10 target nucleic acid. In yet other embodiments, the kit further comprises a third oligonucleotide comprising a 5' portion complementary to a first region of the second target nucleic acid. In some specific embodiments, the 3' portion of the third oligonucleotide is covalently linked to the second target nucleic acid. In other specific

15 embodiments, the second target nucleic acid further comprises a 5' portion, wherein the 5' portion of the second target nucleic acid is the third oligonucleotide. In still other embodiments, the kit further comprises an ARRESTOR molecule (e.g., ARRESTOR oligonucleotide).

The present invention further provides a composition comprising a cleavage structure, the cleavage structure comprising: a) a target nucleic acid, the target nucleic acid having a first region, a second region, a third region and a fourth region, wherein the first region is located adjacent to and downstream from the second region, the second region is located adjacent to and downstream from the third region and the third region is located adjacent to and downstream from the fourth region; b) a first oligonucleotide complementary to the fourth region of the target nucleic acid; c) a second oligonucleotide

20 having a 5' portion and a 3' portion wherein the 5' portion of the second oligonucleotide contains a sequence complementary to the second region of the target nucleic acid and wherein the 3' portion of the second oligonucleotide contains a sequence complementary to the third region of the target nucleic acid; and d) a third oligonucleotide having a 5' portion and a 3' portion wherein the 5' portion of the third oligonucleotide contains a

25 sequence complementary to the first region of the target nucleic acid and wherein the 3'

30 sequence complementary to the first region of the target nucleic acid and wherein the 3'

portion of the third oligonucleotide contains a sequence complementary to the second region of the target nucleic acid.

The present invention is not limited by the length of the four regions of the target nucleic acid. In one embodiment, the first region of the target nucleic acid has a length of 11 to 50 nucleotides. In another embodiment, the second region of the target nucleic acid has a length of one to three nucleotides. In another embodiment, the third region of the target nucleic acid has a length of six to nine nucleotides. In yet another embodiment, the fourth region of the target nucleic acid has a length of 6 to 50 nucleotides.

The invention is not limited by the nature or composition of the of the first, second, third and fourth oligonucleotides; these oligonucleotides may comprise DNA, RNA, PNA and combinations thereof as well as comprise modified nucleotides, universal bases, adducts, etc. Further, one or more of the first, second, third and the fourth oligonucleotides may contain a dideoxynucleotide at the 3' terminus.

In one preferred embodiment, the target nucleic acid is not completely complementary to at least one of the first, the second, the third and the fourth oligonucleotides. In a particularly preferred embodiment, the target nucleic acid is not completely complementary to the second oligonucleotide.

As noted above, the present invention contemplates the use of structure-specific nucleases in detection methods. In one embodiment, the present invention provides a method of detecting the presence of a target nucleic acid molecule by detecting non-target cleavage products comprising: a) providing: i) a cleavage means, ii) a source of target nucleic acid, the target nucleic acid having a first region, a second region, a third region and a fourth region, wherein the first region is located adjacent to and downstream from the second region, the second region is located adjacent to and downstream from the third region and the third region is located adjacent to and downstream from the fourth region; iii) a first oligonucleotide complementary to the fourth region of the target nucleic acid; iv) a second oligonucleotide having a 5' portion and a 3' portion wherein the 5' portion of the second oligonucleotide contains a sequence complementary to the second region of the target nucleic acid and wherein the 3' portion of the second oligonucleotide contains a sequence complementary to the third region of the target nucleic acid; iv) a third oligonucleotide having a 5' and a 3' portion wherein the 5' portion of the third

oligonucleotide contains a sequence complementary to the first region of the target nucleic acid and wherein the 3' portion of the third oligonucleotide contains a sequence complementary to the second region of the target nucleic acid; b) mixing the cleavage means, the target nucleic acid, the first oligonucleotide, the second oligonucleotide and the third oligonucleotide to create a reaction mixture under reaction conditions such that the first oligonucleotide is annealed to the fourth region of the target nucleic acid and wherein at least the 3' portion of the second oligonucleotide is annealed to the target nucleic acid and wherein at least the 5' portion of the third oligonucleotide is annealed to the target nucleic acid so as to create a cleavage structure and wherein cleavage of the cleavage structure occurs to generate non-target cleavage products, each non-target cleavage product having a 3'-hydroxyl group; and c) detecting the non-target cleavage products.

The invention is not limited by the nature of the target nucleic acid. In one embodiment, the target nucleic acid comprises single-stranded DNA. In another embodiment, the target nucleic acid comprises double-stranded DNA and prior to step c), the reaction mixture is treated such that the double-stranded DNA is rendered substantially single-stranded. In another embodiment, the target nucleic acid comprises RNA and the first and second oligonucleotides comprise DNA.

The invention is not limited by the nature of the cleavage means. In one embodiment, the cleavage means is a structure-specific nuclease; particularly preferred structure-specific nucleases are thermostable structure-specific nucleases.

In another preferred embodiment the thermostable structure specific nuclease is a chimerical nuclease.

In an alternative preferred embodiment, the detection of the non-target cleavage products comprises electrophoretic separation of the products of the reaction followed by visualization of the separated non-target cleavage products.

In another preferred embodiment, one or more of the first, second, and third oligonucleotides contain a dideoxynucleotide at the 3' terminus. When dideoxynucleotide-containing oligonucleotides are employed, the detection of the non-target cleavage products preferably comprises: a) incubating the non-target cleavage products with a template-independent polymerase and at least one labeled nucleoside

triphosphate under conditions such that at least one labeled nucleotide is added to the 3'-hydroxyl group of the non-target cleavage products to generate labeled non-target cleavage products; and b) detecting the presence of the labeled non-target cleavage products. The invention is not limited by the nature of the template-independent polymerase employed; in one embodiment, the template-independent polymerase is selected from the group consisting of terminal deoxynucleotidyl transferase (TdT) and poly A polymerase. When TdT or polyA polymerase are employed in the detection step, the second oligonucleotide may contain a 5' end label, the 5' end label being a different label than the label present upon the labeled nucleoside triphosphate. The invention is not limited by the nature of the 5' end label; a wide variety of suitable 5' end labels are known to the art and include biotin, fluorescein, tetrachlorofluorescein, hexachlorofluorescein, Cy3 amidite, Cy5 amidite and digoxigenin.

In another embodiment, detecting the non-target cleavage products comprises: a) incubating the non-target cleavage products with a template-independent polymerase and at least one nucleoside triphosphate under conditions such that at least one nucleotide is added to the 3'-hydroxyl group of the non-target cleavage products to generate tailed non-target cleavage products; and b) detecting the presence of the tailed non-target cleavage products. The invention is not limited by the nature of the template-independent polymerase employed; in one embodiment, the template-independent polymerase is selected from the group consisting of terminal deoxynucleotidyl transferase (TdT) and poly A polymerase. When TdT or polyA polymerases are employed in the detection step, the second oligonucleotide may contain a 5' end label. The invention is not limited by the nature of the 5' end label; a wide variety of suitable 5' end labels are known to the art and include biotin, fluorescein, tetrachlorofluorescein, hexachlorofluorescein, Cy3 amidite, Cy5 amidite and digoxigenin.

In a preferred embodiment, the reaction conditions comprise providing a source of divalent cations; particularly preferred divalent cations are Mn^{2+} and Mg^{2+} ions.

The present invention further provides a method of detecting the presence of a target nucleic acid molecule by detecting non-target cleavage products comprising: a) providing: i) a cleavage means, ii) a source of target nucleic acid, the target nucleic acid having a first region, a second region and a third region, wherein the first region is

located adjacent to and downstream from the second region and wherein the second region is located adjacent to and downstream from the third region; iii) a first oligonucleotide having a 5' and a 3' portion wherein the 5' portion of the first oligonucleotide contains a sequence complementary to the second region of the target nucleic acid and wherein the 3' portion of the first oligonucleotide contains a sequence complementary to the third region of the target nucleic acid; iv) a second oligonucleotide having a length between eleven to fifteen nucleotides and further having a 5' and a 3' portion wherein the 5' portion of the second oligonucleotide contains a sequence complementary to the first region of the target nucleic acid and wherein the 3' portion of the second oligonucleotide contains a sequence complementary to the second region of the target nucleic acid; b) mixing the cleavage means, the target nucleic acid, the first oligonucleotide and the second oligonucleotide to create a reaction mixture under reaction conditions such that at least the 3' portion of the first oligonucleotide is annealed to the target nucleic acid and wherein at least the 5' portion of the second oligonucleotide is annealed to the target nucleic acid so as to create a cleavage structure and wherein cleavage of the cleavage structure occurs to generate non-target cleavage products, each non-target cleavage product having a 3'-hydroxyl group; and c) detecting the non-target cleavage products. In a preferred embodiment the cleavage means is a structure-specific nuclease, preferably a thermostable structure-specific nuclease.

The invention is not limited by the length of the various regions of the target nucleic acid. In a preferred embodiment, the second region of the target nucleic acid has a length between one to five nucleotides. In another preferred embodiment, one or more of the first and the second oligonucleotides contain a dideoxynucleotide at the 3' terminus. When dideoxynucleotide-containing oligonucleotides are employed, the detection of the non-target cleavage products preferably comprises: a) incubating the non-target cleavage products with a template-independent polymerase and at least one labeled nucleoside triphosphate under conditions such that at least one labeled nucleotide is added to the 3'-hydroxyl group of the non-target cleavage products to generate labeled non-target cleavage products; and b) detecting the presence of the labeled non-target cleavage products. The invention is not limited by the nature of the template-independent polymerase employed; in one embodiment, the template-independent polymerase is

selected from the group consisting of terminal deoxynucleotidyl transferase (TdT) and poly A polymerase. When TdT or polyA polymerase is employed in the detection step, the second oligonucleotide may contain a 5' end label, the 5' end label being a different label than the label present upon the labeled nucleoside triphosphate. The invention is not limited by the nature of the 5' end label; a wide variety of suitable 5' end labels are known to the art and include biotin, fluorescein, tetrachlorofluorescein, hexachlorofluorescein, Cy3 amidite, Cy5 amidite and digoxigenin.

In another embodiment, detecting the non-target cleavage products comprises: a) incubating the non-target cleavage products with a template-independent polymerase and at least one nucleoside triphosphate under conditions such that at least one nucleotide is added to the 3'-hydroxyl group of the non-target cleavage products to generate tailed non-target cleavage products; and b) detecting the presence of the tailed non-target cleavage products. The invention is not limited by the nature of the template-independent polymerase employed; in one embodiment, the template-independent polymerase is selected from the group consisting of terminal deoxynucleotidyl transferase (TdT) and poly A polymerase. When TdT or polyA polymerases are employed in the detection step, the second oligonucleotide may contain a 5' end label. The invention is not limited by the nature of the 5' end label; a wide variety of suitable 5' end labels are known to the art and include biotin, fluorescein, tetrachlorofluorescein, hexachlorofluorescein, Cy3 amidite, Cy5 amidite and digoxigenin.

The novel detection methods of the invention may be employed for the detection of target DNAs and RNAs including, but not limited to, target DNAs and RNAs comprising wild type and mutant alleles of genes, including genes from humans or other animals that are or may be associated with disease or cancer. In addition, the methods of the invention may be used for the detection of and/or identification of strains of microorganisms, including bacteria, fungi, protozoa, ciliates and viruses (and in particular for the detection and identification of RNA viruses, such as HCV).

The present invention further provides novel enzymes designed for direct detection, characterization and quantitation of nucleic acids, particularly RNA. The present invention provides enzymes that recognize specific nucleic acid cleavage structures formed on a target RNA sequence and that cleave the nucleic acid cleavage

structure in a site-specific manner to produce non-target cleavage products. The present invention provides enzymes having an improved ability to specifically cleave a DNA member of a complex comprising DNA and RNA nucleic acid strands.

For example, the present invention provides DNA polymerases that are altered in structure relative to the native DNA polymerases, such that they exhibit altered (*e.g.*, improved) performance in detection assays based on the cleavage of a structure comprising nucleic acid (*e.g.*, RNA). In particular, the altered polymerases of the present invention exhibit improved performance in detection assays based on the cleavage of a DNA member of a cleavage structure (*e.g.*, an invasive cleavage structure) that comprises an RNA target strand.

The improved performance in a detection assay may arise from any one of, or a combination of several improved features. For example, in one embodiment, the enzyme of the present invention may have an improved rate of cleavage (k_{cat}) on a specific targeted structure, such that a larger amount of a cleavage product may be produced in a given time span. In another embodiment, the enzyme of the present invention may have a reduced activity or rate in the cleavage of inappropriate or non-specific structures. For example, in certain embodiments of the present invention, one aspect of improvement is that the differential between the detectable amount of cleavage of a specific structure and the detectable amount of cleavage of any alternative structures is increased. As such, it is within the scope of the present invention to provide an enzyme having a reduced rate of cleavage of a specific target structure compared to the rate of the native enzyme, and having a further reduced rate of cleavage of any alternative structures, such that the differential between the detectable amount of cleavage of the specific structure and the detectable amount of cleavage of any alternative structures is increased. However, the present invention is not limited to enzymes that have an improved differential.

In a preferred embodiment, the enzyme of the present invention is a DNA polymerase having an altered nuclease activity as described above, and also having altered synthetic activity, compared to that of any native DNA polymerase from which the enzyme has been derived. It is especially preferred that the DNA polymerase is altered such that it exhibits reduced synthetic activity as well as improved nuclease activity on RNA targets, compared to that of the native DNA polymerase. Enzymes and

genes encoding enzymes having reduced synthetic activity have been described (*See e.g.*, Kaiser *et al.*, J. Biol. Chem., 274:21387 [1999], Lyamichev *et al.*, Prot. Natl. Acad. Sci., 96:6143 [1999], US. Patents 5,541,311, 5,614,402, 5,795,763 and 6,090,606, incorporated herein by reference in their entireties). The present invention contemplates combined modifications, such that the resulting 5' nucleases are without interfering synthetic activity, and have improved performance in RNA detection assays.

The present invention contemplates a DNA sequence encoding a DNA polymerase altered in sequence relative to the native sequence, such that it exhibits altered nuclease activity from that of the native DNA polymerase. For example, in one embodiment, the DNA sequence encodes an enzyme having an improved rate of cleavage (k_{cat}) on a specific targeted structure, such that a larger amount of a cleavage product may be produced in a given time span. In another embodiment, the DNA encodes an enzyme having a reduced activity or rate in the cleavage of inappropriate or non-specific structures. In certain embodiments, one aspect of improvement is that the differential between the detectable amount of cleavage of a specific structure and the detectable amount of cleavage of any alternative structures is increased. It is within the scope of the present invention to provide a DNA encoding an enzyme having a reduced rate of cleavage of a specific target structure compared to the rate of the native enzyme, and having a further reduced rate of cleavage of any alternative structures, such that the differential between the detectable amount of cleavage of the specific structure and the detectable amount of cleavage of any alternative structures is increased. However, the present invention is not limited to polymerases that have an improved differential.

In a preferred embodiment, the DNA sequence encodes a DNA polymerase having the altered nuclease activity described above, and also having altered synthetic activity, compared to that of any native DNA polymerase from which the improved enzyme is derived. It is especially preferred that the encoded DNA polymerase is altered such that it exhibits reduced synthetic activity as well as improved nuclease activity on RNA targets, compared to that of the native DNA polymerase.

It is not intended that the invention be limited by the nature of the alteration required to introduce altered nuclease activity. Nor is it intended that the invention be limited by the extent of either the alteration, or in the improvement observed. If the

polymerase is also altered so as to be synthesis modified, it is not intended that the invention be limited by the polymerase activity of the modified or unmodified protein, or by the nature of the alteration to render the polymerase synthesis modified.

The present invention contemplates structure-specific nucleases from a variety of sources, including, but not limited to, mesophilic, psychrophilic, thermophilic, and hyperthermophilic organisms. The preferred structure-specific nucleases are thermostable. Thermostable structure-specific nucleases are contemplated as particularly useful in that they allow the INVADER assay (*See e.g.*, U.S. Pat. Nos. 5,846,717, 5,985,557, 5,994,069, 6,001,567, and 6,090,543 and PCT Publications WO 97/27214 and WO 98/42873, incorporated herein by reference in their entireties) to be operated near the melting temperature (T_m) of the downstream probe oligonucleotide, so that cleaved and uncleaved probes may cycle on and off the target during the course of the reaction. In one embodiment, the thermostable structure-specific enzymes are thermostable 5' nucleases that are selected from the group comprising altered polymerases derived from the native polymerases of *Thermus* species, including, but not limited to, *Thermus aquaticus*, *Thermus flavus*, *Thermus thermophilus*, *Thermus filiformus*, and *Thermus scotoductus*. However, the invention is not limited to the use of thermostable 5' nucleases. For example, certain embodiments of the present invention utilize short oligonucleotide probes that may cycle on and off of the target at low temperatures, allowing the use of non-thermostable enzymes.

In some preferred embodiments, the present invention provides a composition comprising an enzyme, wherein the enzyme comprises a heterologous functional domain, wherein the heterologous functional domain provides altered (*e.g.*, improved) functionality in a nucleic acid cleavage assay. The present invention is not limited by the nature of the nucleic acid cleavage assay. For example, nucleic acid cleavage assays include any assay in which a nucleic acid is cleaved, directly or indirectly, in the presence of the enzyme. In certain preferred embodiments, the nucleic acid cleavage assay is an invasive cleavage assay. In particularly preferred embodiments, the cleavage assay utilizes a cleavage structure having at least one RNA component. In another particularly preferred embodiment, the cleavage assay utilizes a cleavage structure having at least one RNA component, wherein a DNA member of the cleavage structure is cleaved.

The present invention is not limited by the nature of the altered functionality provided by the heterologous functional domain. Illustrative examples of alterations include, but are not limited to, enzymes where the heterologous functional domain comprises an amino acid sequence (*e.g.*, one or more amino acids) that provides an improved nuclease activity, an improved substrate binding activity and/or improved background specificity in a nucleic acid cleavage assay.

The present invention is not limited by the nature of the heterologous functional domain. For example, in some embodiments, the heterologous functional domain comprises two or more amino acids from a polymerase domain of a polymerase (e.g., introduced into the enzyme by insertion of a chimeric functional domain or created by mutation). In certain preferred embodiment, at least one of the two or more amino acids is from a palm or thumb region of the polymerase domain. The present invention is not limited by the identity of the polymerase from which the two or more amino acids are selected. In certain preferred embodiments, the polymerase comprises *Thermus thermophilus* polymerase. In particularly preferred embodiments, the two or more amino acids are from amino acids 300-650 of SEQ ID NO:1.

The novel enzymes of the invention may be employed for the detection of target DNAs and RNAs including, but not limited to, target DNAs and RNAs comprising wild type and mutant alleles of genes, including, but not limited to, genes from humans, other animal, or plants that are or may be associated with disease or other conditions. In addition, the enzymes of the invention may be used for the detection of and/or identification of strains of microorganisms, including bacteria, fungi, protozoa, ciliates and viruses (and in particular for the detection and identification of viruses having RNA

genomes, such as the Hepatitis C and Human Immunodeficiency viruses). For example, the present invention provides methods for cleaving a nucleic acid comprising providing: an enzyme of the present invention and a substrate nucleic acid; and exposing the substrate nucleic acid to the enzyme (*e.g.*, to produce a cleavage product that may be detected).

In one embodiment, the present invention provides a thermostable 5' nuclease having an amino acid sequence selected from the group comprising SEQ ID NOS:2, 3, 4, 5, 6, 7, 8, 9, 10, 11 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 341, 346, 348, 351, 353, 359, 365, 367, 369, 374, 376, 380, 384, 388, 392, 396, 400, 402, 406, 408, 410, 412, 416, 418, 420, 424, 427, 429, 432, 436, 440, 444, 446, 448, 450, 456, 460, 464, 468, 472, 476, 482, 485, 488, 491, 494, 496, 498, 500, 502, 506, 510, 514, 518, 522, 526, 530, 534, 538, 542, 544, 550, 553, 560, 564, 566, 568, 572, 574, 576, 578, 580, 582, 584, 586, 588, and 590. In another embodiment, the 5' nuclease is encoded by a DNA sequence selected from the group comprising of SEQ ID NOS:69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 340, 345, 347, 350, 352, 358, 364, 366, 368, 373, 375, 379, 383, 387, 391, 395, 399, 401, 405, 407, 409, 411, 415, 417, 419, 423, 426, 428, 431, 435, 439, 443, 445, 447, 449, 452, 454, 455, 459, 463, 467, 471, 475, 481, 484, 495, 497, 499, 501, 505, 509, 513, 517, 521, 525, 529, 533, 537, 541, 543, 549, 552, 559, 563, 565, 567, 571, 573, 575, 577, 579, 581, 583, 585, 587, and 589.

The present invention also provides a recombinant DNA vector comprising DNA having a nucleotide sequence encoding a 5' nuclease, the nucleotide sequence selected from the group comprising SEQ ID NOS: 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 340, 345, 347, 350, 352, 358, 364, 366, 368, 373, 375, 379, 383, 387, 391, 395, 399, 401, 405, 407, 409, 411, 415, 417, 419, 423, 426, 428, 431, 435, 439, 443, 445, 447, 449, 452, 454, 455, 459,

463, 467, 471, 475, 481, 484, 495, 497, 499, 501, 505, 509, 513, 517, 521, 525, 529, 533, 537, 541, 543, 549, 552, 559, 563, 565, 567, 571, 573, 575, 577, 579, 581, 583, 585, 587, and 589. In a preferred embodiment, the invention provides a host cell transformed with a recombinant DNA vector comprising DNA having a nucleotide sequence encoding a

5 structure-specific nuclease, the nucleotide selected from the group comprising sequence SEQ ID NOS: 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 340, 345, 347, 350, 352, 358, 364, 366, 368, 10 373, 375, 379, 383, 387, 391, 395, 399, 401, 405, 407, 409, 411, 415, 417, 419, 423, 426, 428, 431, 435, 439, 443, 445, 447, 449, 452, 454, 455, 459, 463, 467, 471, 475, 481, 484, 495, 497, 499, 501, 505, 509, 513, 517, 521, 525, 529, 533, 537, 541, 543, 549, 552, 559, 563, 565, 567, 571, 573, 575, 577, 579, 581, 583, 585, 587, and 589. The invention is not limited by the nature of the host cell employed. The art is well aware of expression 15 vectors suitable for the expression of nucleotide sequences encoding structure-specific nucleases that can be expressed in a variety of prokaryotic and eukaryotic host cells. In a preferred embodiment, the host cell is an *Escherichia coli* cell.

The present invention provides a method of altering 5' nuclease enzymes relative to native 5' nuclease enzymes, such that they exhibit improved performance in detection 20 assays based on the cleavage of a structure comprising RNA. In particular, the altered 5' nucleases produced by the method of the present invention exhibit improved performance in detection assays based on the cleavage of a DNA member of a cleavage structure (*e.g.*, an invasive cleavage structure) that comprises an RNA target strand. The improved 5' nucleases resulting from the methods of the present invention may be improved in any of 25 the ways discussed herein. Examples of processes for assessing improvement in any candidate enzyme are provided.

For example, the present invention provides methods for producing an altered enzyme with improved functionality in a nucleic acid cleavage assay comprising: providing an enzyme and a nucleic acid test substrate; introducing a heterologous 30 functional domain into the enzyme to produce an altered enzyme; contacting the altered enzyme with the nucleic acid test substrate to produce cleavage products; and detecting

the cleavage products. In some embodiments, the introduction of the heterologous functional domain comprises mutating one or more amino acids of the enzyme. In other embodiments, the introduction of the heterologous functional domain into the enzyme comprises adding a functional domain from a protein (e.g., another enzyme) into the enzyme (e.g., substituting functional domains by removing a portion of the enzyme sequence prior to adding the functional domain of the protein). In preferred embodiments, the nucleic acid test substrate comprises a cleavage structure. In particularly preferred embodiment, the cleavage structure comprises an RNA target nucleic acid. In yet other preferred embodiments, the cleavage structure comprises an invasive cleavage structure.

The present invention also provides nucleic acid treatment kits. One preferred embodiment is a kit comprising a composition comprising at least one improved 5' nuclease. Another preferred embodiment provides a kit comprising: a) a composition comprising at least one improved 5' nuclease; and b) an INVADER oligonucleotide and a signal probe oligonucleotide. In some embodiments of the kits of the present invention, the improved 5' nuclease is derived from a DNA polymerase from a eubacterial species. In further embodiments, the eubacterial species is a thermophile. In still further embodiments, the thermophile is of the genus *Thermus*. In still further embodiments, the thermophile is selected from the group consisting of *Thermus aquaticus*, *Thermus flavus*, *Thermus thermophilus*, *Thermus filiformus*, and *Thermus scotoductus*. In preferred embodiments, the improved 5' nuclease is encoded by DNA selected from the group comprising SEQ ID NOS: 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 340, 345, 347, 350, 352, 358, 364, 366, 368, 373, 375, 379, 383, 387, 391, 395, 399, 401, 405, 407, 409, 411, 415, 417, 419, 423, 426, 428, 431, 435, 439, 443, 445, 447, 449, 452, 454, 455, 459, 463, 467, 471, 475, 481, 484, 495, 497, 499, 501, 505, 509, 513, 517, 521, 525, 529, 533, 537, 541, 543, 549, 552, 559, 563, 565, 567, 571, 573, 575, 577, 579, 581, 583, 585, 587, and 589. In yet other preferred embodiments, the kits further comprise reagents for detecting a nucleic acid cleavage product. In further preferred embodiments, the reagents for detecting a

cleavage product comprise oligonucleotides for use in a subsequent invasive cleavage reaction (See e.g., U.S. Patent No. 5,994,069). In particularly preferred embodiments, the reagents for the subsequent invasive cleavage reaction comprise a probe labeled with moieties that produce a fluorescence resonance energy transfer (FRET) effect.

5 The present invention also provides methods for treating nucleic acid, comprising:
 a) providing: a first structure-specific nuclease consisting of an endonuclease in a solution containing manganese; and a nucleic acid substrate; b) treating the nucleic acid substrate with increased temperature such that the substrate is substantially single-stranded; c) reducing the temperature under conditions such that the single-stranded
 10 substrate forms one or more cleavage structures; d) reacting the cleavage means with the cleavage structures so that one or more cleavage products are produced; and e) detecting the one or more cleavage products. In some embodiments of the methods, the endonuclease includes, but is not limited to, CLEAVASE BN enzyme, *Thermus aquaticus* DNA polymerase, *Thermus thermophilus* DNA polymerase, *Escherichia coli*
 15 Exo III, and the *Saccharomyces cerevisiae* Rad1/Rad10 complex. In yet other preferred embodiments, the nuclease is a 5' nuclease derived from a thermostable DNA polymerase altered in amino acid sequence such that it exhibits reduced DNA synthetic activity from that of the wild-type DNA polymerase but retains substantially the same 5' nuclease activity of the wild-type DNA polymerase. In yet other embodiments, the nucleic acid is
 20 selected from the group consisting of RNA and DNA. In further embodiments, the nucleic acid of step (a) is double stranded.

The present invention also provides nucleic acid treatment kits, comprising: a) a composition comprising at least one purified FEN-1 endonuclease; and b) a solution containing manganese. In some embodiments of the kits, the purified FEN-1
 25 endonuclease is selected from the group consisting *Pyrococcus woesei* FEN-1 endonuclease, *Pyrococcus furiosus* FEN-1 endonuclease, *Methanococcus jannaschii* FEN-1 endonuclease, *Methanobacterium thermoautotrophicum* FEN-1 endonuclease, *Archaeoglobus fulgidus* FEN-1, *Sulfolobus solfataricus*, *Pyrobaculum aerophilum*, *Thermococcus litoralis*, *Archaeoglobus veneficus*, *Archaeoglobus profundus*, *Acidianus*
 30 *brierlyi*, *Acidianus ambivalens*, *Desulfurococcus amylolyticus*, *Desulfurococcus mobilis*, *Pyrodictium brockii*, *Thermococcus gorgonarius*, *Thermococcus zilligii*, *Methanopyrus*

kandleri, *Methanococcus igneus*, *Pyrococcus horikoshii*, *Aeropyrum pernix*, and chimerical FEN-1 endonucleases. In other embodiments, the kits further comprise at least one second structure-specific nuclease. In some preferred embodiments, the second nuclease is a 5' nuclease derived from a thermostable DNA polymerase altered in amino acid sequence such that it exhibits reduced DNA synthetic activity from that of the wild-type DNA polymerase but retains substantially the same 5' nuclease activity of the wild-type DNA polymerase. In yet other embodiments of the kits, the portion of the amino acid sequence of the second nuclease is homologous to a portion of the amino acid sequence of a thermostable DNA polymerase derived from a eubacterial thermophile of the genus *Thermus*. In further embodiments, the thermophile is selected from the group consisting of *Thermus aquaticus*, *Thermus flavus* and *Thermus thermophilus*. In yet other preferred embodiments, the kits further comprise reagents for detecting the cleavage products.

The present invention further provides any of the compositions, mixtures, methods, and kits described herein, used in conjunction with endonucleases comprising *Sulfolobus solfataricus*, *Pyrobaculum aerophilum*, *Thermococcus litoralis*, *Archaeoglobus veneficus*, *Archaeoglobus profundus*, *Acidianus brierlyi*, *Acidianus ambivalens*, *Desulfurococcus amylolyticus*, *Desulfurococcus mobilis*, *Pyrodictium brockii*, *Thermococcus gorgonarius*, *Thermococcus zilligii*, *Methanopyrus kandleri*, *Methanococcus igneus*, *Pyrococcus horikoshii*, and *Aeropyrum pernix* endonucleases. These include compositions comprising purified FEN-1 endonucleases from the organisms (including specific endonucleases described by sequences provided herein, as well as, variants and homologues), kits comprising these compositions, composition comprising chimerical endonucleases comprising at least a portion of the endonucleases from these organisms, kits comprising such compositions, compositions comprising nucleic acids encoding the endonucleases from these organisms (including vectors and host cells), kits comprising such compositions, antibodies generated to the endonucleases, mixtures comprising endonucleases from these organisms, methods of using the endonuclease in cleavage assays (e.g., invasive cleavage assays, CFLP, etc.), and kits containing components useful for such methods. Examples describing the generation, structure, use, and characterization of these endonucleases are provided herein.

The present invention also provides methods for improving the methods and enzymes disclosed herein. For example, the present invention provides methods of improving enzymes for any intended purpose (e.g., use in cleavage reactions, amplification reactions, binding reactions, or any other use) comprising the step of providing an enzyme disclosed herein and modifying the enzyme (e.g., altering the amino acid sequence, adding or subtracting sequence, adding post-translational modifications, adding any other component whether biological or not, or any other modification). Likewise, the present invention provides methods for improving the methods disclosed herein comprising, conducting the method steps with one or more changes (e.g., change in a composition provided in the method, change in the order of the steps, or addition or subtraction of steps).

The improved performance in a detection assay may arise from any one of, or a combination of several improved features. For example, in one embodiment, the enzyme of the present invention may have an improved rate of cleavage (k_{cat}) on a specific targeted structure, such that a larger amount of a cleavage product may be produced in a given time span. In another embodiment, the enzyme of the present invention may have a reduced activity or rate in the cleavage of inappropriate or non-specific structures. For example, in certain embodiments of the present invention, one aspect of improvement is that the differential between the detectable amount of cleavage of a specific structure and the detectable amount of cleavage of any alternative structures is increased. As such, it is within the scope of the present invention to provide an enzyme having a reduced rate of cleavage of a specific target structure compared to the rate of the native enzyme, and having a further reduced rate of cleavage of any alternative structures, such that the differential between the detectable amount of cleavage of the specific structure and the detectable amount of cleavage of any alternative structures is increased. However, the present invention is not limited to enzymes that have an improved differential.

In some preferred embodiments, the present invention provides a composition comprising an enzyme, wherein the enzyme comprises a heterologous functional domain, wherein the heterologous functional domain provides altered (e.g., improved) functionality in a nucleic acid cleavage assay. The present invention is not limited by the nature of the nucleic acid cleavage assay. For example, nucleic acid cleavage assays

include any assay in which a nucleic acid is cleaved, directly or indirectly, in the presence of the enzyme. In certain preferred embodiments, the nucleic acid cleavage assay is an invasive cleavage assay. In particularly preferred embodiments, the cleavage assay utilizes a cleavage structure having at least one RNA component. In another particularly preferred embodiment, the cleavage assay utilizes a cleavage structure having at least one RNA component, wherein a DNA member of the cleavage structure is cleaved.

The present invention is not limited by the nature of the altered functionality provided by the heterologous functional domain. Illustrative examples of alterations include, but are not limited to, enzymes where the heterologous functional domain comprises an amino acid sequence (e.g., one or more amino acids) that provides an improved nuclease activity, an improved substrate binding activity and/or improved background specificity in a nucleic acid cleavage assay.

The present invention is not limited by the nature of the heterologous functional domain. For example, in some embodiments, the heterologous functional domain comprises two or more amino acids from a polymerase domain of a polymerase (e.g., introduced into the enzyme by insertion of a chimerical functional domain or created by mutation). In certain preferred embodiment, at least one of the two or more amino acids is from a palm or thumb region of the polymerase domain. The present invention is not limited by the identity of the polymerase from which the two or more amino acids are selected. In certain preferred embodiments, the polymerase comprises *Thermus thermophilus* polymerase. In particularly preferred embodiments, the two or more amino acids are from amino acids 300-650 of SEQ ID NO:1.

The novel enzymes of the invention may be employed for the detection of target DNAs and RNAs including, but not limited to, target DNAs and RNAs comprising wild type and mutant alleles of genes, including, but not limited to, genes from humans, other animal, or plants that are or may be associated with disease or other conditions. In addition, the enzymes of the invention may be used for the detection of and/or identification of strains of microorganisms, including bacteria, fungi, protozoa, ciliates and viruses (and in particular for the detection and identification of viruses having RNA genomes, such as the Hepatitis C and Human Immunodeficiency viruses). For example, the present invention provides methods for cleaving a nucleic acid comprising providing:

an enzyme of the present invention and a substrate nucleic acid; and exposing the substrate nucleic acid to the enzyme (e.g., to produce a cleavage product that may be detected). In some embodiments, the substrate nucleic is in a cell lysate sample.

The present invention also provides a method for detecting the presence of a target nucleic acid comprising: cleaving an invasive cleavage structure, said invasive cleavage structure comprising an RNA target nucleic acid; and detecting the cleavage of the invasive cleavage structure. Such an assay may comprise a multiplex assay, wherein multiple invasive cleavage structures are cleaved. Such structures include structures formed on different target nucleic acids, as well as, structures formed on different locations of the sample target nucleic acid. In some embodiments, the target nucleic acid comprises a first region and a second region, said second region downstream of and contiguous to said first region. In some embodiments, the invasive cleavage structure comprises the target nucleic acid, a first oligonucleotide, and a second oligonucleotide, wherein at least a portion of the first oligonucleotide is completely complementary to the first portion of the first target nucleic acid, and wherein the second oligonucleotide comprises a 3' portion and a 5' portion, wherein the 5' portion is completely complementary to said second portion of the target nucleic acid. In some embodiments, the 3' portion of the second oligonucleotide comprises a 3' terminal nucleotide not complementary to said target nucleic acid. In some embodiments, the 3' portion of the second oligonucleotide consists of a single nucleotide not complementary to the target nucleic acid. In some embodiments, the method further comprises the steps of forming a second invasive cleavage structure comprising a non-target cleavage product and cleaving the second invasive cleavage structure. In some embodiments, the invasive cleavage structure or the second invasive cleavage comprises an oligonucleotide comprising a sequence selected from the group consisting of SEQ ID NO:709-2640. In other embodiments, the invasive cleavage structure or the second invasive cleavage comprises an oligonucleotide comprising a sequence selected from the group consisting of SEQ ID NO:169-211 and 619-706. In some preferred embodiments, the target nucleic acid comprises a cytochrome P450 RNA or a cytokine RNA. In some embodiments, the first region or the second region of the target nucleic acid encompasses a splice junction, an exon (or a portion thereof), or an intron (or a portion thereof). In some embodiments,

the RNA target nucleic acid is provided in a cell lysate. In some embodiments, the first oligonucleotide is covalently attached to the second oligonucleotide. Such oligonucleotides find use, for example, in methods described in U.S. Patent Nos. 5,714,320 and 5,854,033, herein incorporated by reference in their entireties. The present invention also provides kits containing one or more of the components used in the above methods.

DEFINITIONS

To facilitate an understanding of the present invention, a number of terms and phrases are defined below:

As used herein, the terms "complementary" or "complementarity" are used in reference to polynucleotides (*i.e.*, a sequence of nucleotides such as an oligonucleotide or a target nucleic acid) related by the base-pairing rules. For example, for the sequence "5'-A-G-T-3'," is complementary to the sequence "3'-T-C-A-5'." Complementarity may be "partial," in which only some of the nucleic acids' bases are matched according to the base pairing rules. Or, there may be "complete" or "total" complementarity between the nucleic acids. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of hybridization between nucleic acid strands. This is of particular importance in amplification reactions, as well as detection methods that depend upon binding between nucleic acids. Either term may also be used in reference to individual nucleotides, especially within the context of polynucleotides. For example, a particular nucleotide within an oligonucleotide may be noted for its complementarity, or lack thereof, to a nucleotide within another nucleic acid strand, in contrast or comparison to the complementarity between the rest of the oligonucleotide and the nucleic acid strand. Nucleotide analogs used to form non-standard base pairs, whether with another nucleotide analog (e.g., an IsoC/IsoG base pair), or with a naturally occurring nucleotide (e.g., as described in U.S. Patent 5,912,340, herein incorporated by reference in its entirety) are also considered to be complementary to a base pairing partner within the meaning this definition.

The term "homology" and "homologous" refers to a degree of identity. There may be partial homology or complete homology. A partially homologous sequence is one that is less than 100% identical to another sequence.

As used herein, the term "hybridization" is used in reference to the pairing of complementary nucleic acids. Hybridization and the strength of hybridization (*i.e.*, the strength of the association between the nucleic acids) is influenced by such factors as the degree of complementarity between the nucleic acids, stringency of the conditions involved, and the T_m of the formed hybrid. "Hybridization" methods involve the annealing of one nucleic acid to another, complementary nucleic acid, *i.e.*, a nucleic acid having a complementary nucleotide sequence. The ability of two polymers of nucleic acid containing complementary sequences to find each other and anneal through base pairing interaction is a well-recognized phenomenon. The initial observations of the "hybridization" process by Marmur and Lane, Proc. Natl. Acad. Sci. USA 46:453 (1960) and Doty et al., Proc. Natl. Acad. Sci. USA 46:461 (1960) have been followed by the refinement of this process into an essential tool of modern biology.

With regard to complementarity, it is important for some diagnostic applications to determine whether the hybridization represents complete or partial complementarity. For example, where it is desired to detect simply the presence or absence of pathogen DNA (such as from a virus, bacterium, fungi, mycoplasma, protozoan) it is only important that the hybridization method ensures hybridization when the relevant sequence is present; conditions can be selected where both partially complementary probes and completely complementary probes will hybridize. Other diagnostic applications, however, may require that the hybridization method distinguish between partial and complete complementarity. It may be of interest to detect genetic polymorphisms. For example, human hemoglobin is composed, in part, of four polypeptide chains. Two of these chains are identical chains of 141 amino acids (alpha chains) and two of these chains are identical chains of 146 amino acids (beta chains). The gene encoding the beta chain is known to exhibit polymorphism. The normal allele encodes a beta chain having glutamic acid at the sixth position. The mutant allele encodes a beta chain having valine at the sixth position. This difference in amino acids has a profound (most profound when the individual is homozygous for the mutant allele)

physiological impact known clinically as sickle cell anemia. It is well known that the genetic basis of the amino acid change involves a single base difference between the normal allele DNA sequence and the mutant allele DNA sequence.

The complement of a nucleic acid sequence as used herein refers to an oligonucleotide which, when aligned with the nucleic acid sequence such that the 5' end of one sequence is paired with the 3' end of the other, is in "antiparallel association." Certain bases not commonly found in natural nucleic acids may be included in the nucleic acids of the present invention and include, for example, inosine and 7-deazaguanine. Complementarity need not be perfect; stable duplexes may contain mismatched base pairs or unmatched bases. Those skilled in the art of nucleic acid technology can determine duplex stability empirically considering a number of variables including, for example, the length of the oligonucleotide, base composition and sequence of the oligonucleotide, ionic strength and incidence of mismatched base pairs.

As used herein, the term " T_m " is used in reference to the "melting temperature." The melting temperature is the temperature at which a population of double-stranded nucleic acid molecules becomes half dissociated into single strands. Several equations for calculating the T_m of nucleic acids are well known in the art. As indicated by standard references, a simple estimate of the T_m value may be calculated by the equation: $T_m = 81.5 + 0.41(\% G + C)$, when a nucleic acid is in aqueous solution at 1 M NaCl (*see e.g.*, Anderson and Young, Quantitative Filter Hybridization, in Nucleic Acid Hybridization (1985). Other references (*e.g.*, Allawi, H.T. & SantaLucia, J., Jr. Thermodynamics and NMR of internal G.T mismatches in DNA. Biochemistry 36, 10581-94 (1997) include more sophisticated computations which take structural and environmental, as well as sequence characteristics into account for the calculation of T_m .

As used herein the term "stringency" is used in reference to the conditions of temperature, ionic strength, and the presence of other compounds, under which nucleic acid hybridizations are conducted. With "high stringency" conditions, nucleic acid base pairing will occur only between nucleic acid fragments that have a high frequency of complementary base sequences. Thus, conditions of "weak" or "low" stringency are often required when it is desired that nucleic acids that are not completely complementary to one another be hybridized or annealed together.

"High stringency conditions" when used in reference to nucleic acid hybridization comprise conditions equivalent to binding or hybridization at 42 C in a solution consisting of 5X SSPE (43.8 g/l NaCl, 6.9 g/l NaH₂PO₄ H₂O and 1.85 g/l EDTA, pH adjusted to 7.4 with NaOH), 0.5% SDS, 5X Denhardt's reagent and 100 µg/ml denatured salmon sperm DNA followed by washing in a solution comprising 0.1X SSPE, 1.0% SDS at 42 C when a probe of about 500 nucleotides in length is employed.

"Medium stringency conditions" when used in reference to nucleic acid hybridization comprise conditions equivalent to binding or hybridization at 42 C in a solution consisting of 5X SSPE (43.8 g/l NaCl, 6.9 g/l NaH₂PO₄ H₂O and 1.85 g/l EDTA, pH adjusted to 7.4 with NaOH), 0.5% SDS, 5X Denhardt's reagent and 100 µg/ml denatured salmon sperm DNA followed by washing in a solution comprising 1.0X SSPE, 1.0% SDS at 42 C when a probe of about 500 nucleotides in length is employed.

"Low stringency conditions" comprise conditions equivalent to binding or hybridization at 42 C in a solution consisting of 5X SSPE (43.8 g/l NaCl, 6.9 g/l NaH₂PO₄ H₂O and 1.85 g/l EDTA, pH adjusted to 7.4 with NaOH), 0.1% SDS, 5X Denhardt's reagent [50X Denhardt's contains per 500 ml: 5 g Ficoll (Type 400, Pharmacia), 5 g BSA (Fraction V; Sigma)] and 100 g/ml denatured salmon sperm DNA followed by washing in a solution comprising 5X SSPE, 0.1% SDS at 42 C when a probe of about 500 nucleotides in length is employed.

The term "gene" refers to a DNA sequence that comprises control and coding sequences necessary for the production of an RNA having a non-coding function (*e.g.*, a ribosomal or transfer RNA), a polypeptide or a precursor. The RNA or polypeptide can be encoded by a full-length coding sequence or by any portion of the coding sequence so long as the desired activity or function is retained.

The term "wild-type" refers to a gene or a gene product that has the characteristics of that gene or gene product when isolated from a naturally occurring source. A wild-type gene is that which is most frequently observed in a population and is thus arbitrarily designated the "normal" or "wild-type" form of the gene. In contrast, the term "modified," "mutant," or "polymorphic" refers to a gene or gene product that displays modifications in sequence and or functional properties (*i.e.*, altered characteristics) when compared to the wild-type gene or gene product. It is noted that naturally-occurring

mutants can be isolated; these are identified by the fact that they have altered characteristics when compared to the wild-type gene or gene product.

The term "recombinant DNA vector" as used herein refers to DNA sequences containing a desired heterologous sequence. For example, although the term is not limited to the use of expressed sequences or sequences that encode an expression product, in some embodiments, the heterologous sequence is a coding sequence and appropriate DNA sequences necessary for either the replication of the coding sequence in a host organism, or the expression of the operably linked coding sequence in a particular host organism. DNA sequences necessary for expression in prokaryotes include a promoter, optionally an operator sequence, a ribosome-binding site and possibly other sequences. Eukaryotic cells are known to utilize promoters, polyadenylation signals and enhancers.

The term "LTR" as used herein refers to the long terminal repeat found at each end of a provirus (*i.e.*, the integrated form of a retrovirus). The LTR contains numerous regulatory signals including transcriptional control elements, polyadenylation signals and sequences needed for replication and integration of the viral genome. The viral LTR is divided into three regions called U3, R and U5.

The U3 region contains the enhancer and promoter elements. The U5 region contains the polyadenylation signals. The R (repeat) region separates the U3 and U5 regions and transcribed sequences of the R region appear at both the 5' and 3' ends of the viral RNA.

The term "oligonucleotide" as used herein is defined as a molecule comprising two or more deoxyribonucleotides or ribonucleotides, preferably at least 5 nucleotides, more preferably at least about 10-15 nucleotides and more preferably at least about 15 to 30 nucleotides. The exact size will depend on many factors, which in turn depend on the ultimate function or use of the oligonucleotide. The oligonucleotide may be generated in any manner, including chemical synthesis, DNA replication, reverse transcription, PCR, or a combination thereof.

Because mononucleotides are reacted to make oligonucleotides in a manner such that the 5' phosphate of one mononucleotide pentose ring is attached to the 3' oxygen of its neighbor in one direction via a phosphodiester linkage, an end of an oligonucleotide is referred to as the "5' end" if its 5' phosphate is not linked to the 3' oxygen of a

mononucleotide pentose ring and as the "3' end" if its 3' oxygen is not linked to a 5' phosphate of a subsequent mononucleotide pentose ring. As used herein, a nucleic acid sequence, even if internal to a larger oligonucleotide, also may be said to have 5' and 3' ends. A first region along a nucleic acid strand is said to be upstream of another region if the 3' end of the first region is before the 5' end of the second region when moving along a strand of nucleic acid in a 5' to 3' direction.

When two different, non-overlapping oligonucleotides anneal to different regions of the same linear complementary nucleic acid sequence, and the 3' end of one oligonucleotide points towards the 5' end of the other, the former may be called the "upstream" oligonucleotide and the latter the "downstream" oligonucleotide. Similarly, when two overlapping oligonucleotides are hybridized to the same linear complementary nucleic acid sequence, with the first oligonucleotide positioned such that its 5' end is upstream of the 5' end of the second oligonucleotide, and the 3' end of the first oligonucleotide is upstream of the 3' end of the second oligonucleotide, the first oligonucleotide may be called the "upstream" oligonucleotide and the second oligonucleotide may be called the "downstream" oligonucleotide.

The term "primer" refers to an oligonucleotide that is capable of acting as a point of initiation of synthesis when placed under conditions in which primer extension is initiated. An oligonucleotide "primer" may occur naturally, as in a purified restriction digest or may be produced synthetically.

A primer is selected to be "substantially" complementary to a strand of specific sequence of the template. A primer must be sufficiently complementary to hybridize with a template strand for primer elongation to occur. A primer sequence need not reflect the exact sequence of the template. For example, a non-complementary nucleotide fragment may be attached to the 5' end of the primer, with the remainder of the primer sequence being substantially complementary to the strand. Non-complementary bases or longer sequences can be interspersed into the primer, provided that the primer sequence has sufficient complementarity with the sequence of the template to hybridize and thereby form a template primer complex for synthesis of the extension product of the primer.

1 The term "label" as used herein refers to any atom or molecule that can be used to
provide a detectable (preferably quantifiable) effect, and that can be attached to a nucleic
acid or protein. Labels include but are not limited to dyes; radiolabels such as ³²P;
binding moieties such as biotin; haptens such as digoxigenin; luminogenic,
5 phosphorescent or fluorogenic moieties; and fluorescent dyes alone or in combination
with moieties that can suppress or shift emission spectra by fluorescence resonance
energy transfer (FRET). Labels may provide signals detectable by fluorescence,
radioactivity, colorimetry, gravimetry, X-ray diffraction or absorption, magnetism,
enzymatic activity, and the like. A label may be a charged moiety (positive or negative
10 charge) or alternatively, may be charge neutral. Labels can include or consist of nucleic
acid or protein sequence, so long as the sequence comprising the label is detectable.

The term "signal" as used herein refers to any detectable effect, such as would be
caused or provided by a label or an assay reaction.

As used herein, the term "detector" refers to a system or component of a system,
15 e.g., an instrument (e.g. a camera, fluorimeter, charge-coupled device, scintillation
counter, etc.) or a reactive medium (X-ray or camera film, pH indicator, etc.), that can
convey to a user or to another component of a system (e.g., a computer or controller) the
presence of a signal or effect. A detector can be a photometric or spectrophotometric
system, which can detect ultraviolet, visible or infrared light, including fluorescence or
20 chemiluminescence; a radiation detection system; a spectroscopic system such as nuclear
magnetic resonance spectroscopy, mass spectrometry or surface enhanced Raman
spectrometry; a system such as gel or capillary electrophoresis or gel exclusion
chromatography; or other detection systems known in the art, or combinations thereof.

The term "cleavage structure" as used herein, refers to a structure that is formed
25 by the interaction of at least one probe oligonucleotide and a target nucleic acid, forming
a structure comprising a duplex, the resulting structure being cleavable by a cleavage
agent, including but not limited to an enzyme. The cleavage structure is a substrate for
specific cleavage by the cleavage means in contrast to a nucleic acid molecule that is a
substrate for non-specific cleavage by agents such as phosphodiesterases that cleave
30 nucleic acid molecules without regard to secondary structure (*i.e.*, no formation of a
duplexed structure is required).

The term "folded cleavage structure" as used herein, refers to a region of a single-stranded nucleic acid substrate containing secondary structure, the region being cleavable by an enzymatic cleavage means. The cleavage structure is a substrate for specific cleavage by the cleavage means in contrast to a nucleic acid molecule that is a substrate for non-specific cleavage by agents such as phosphodiesterases that cleave nucleic acid molecules without regard to secondary structure (*i.e.*, no folding of the substrate is required).

As used herein, the term "folded target" refers to a nucleic acid strand that contains at least one region of secondary structure (*i.e.*, at least one double stranded region and at least one single-stranded region within a single strand of the nucleic acid). A folded target may comprise regions of tertiary structure in addition to regions of secondary structure.

The term "cleavage means" or "cleavage agent" as used herein refers to any means that is capable of cleaving a cleavage structure, including but not limited to enzymes. The cleavage means may include native DNAPs having 5' nuclease activity (*e.g.*, Taq DNA polymerase, *E. coli* DNA polymerase I) and, more specifically, modified DNAPs having 5' nuclease but lacking synthetic activity. "Structure-specific nucleases" or "structure-specific enzymes" are enzymes that recognize specific secondary structures in a nucleic acid molecule and cleave these structures. The cleavage means of the invention cleave a nucleic acid molecule in response to the formation of cleavage structures; it is not necessary that the cleavage means cleave the cleavage structure at any particular location within the cleavage structure.

The cleavage means is not restricted to enzymes having solely 5' nuclease activity. The cleavage means may include nuclease activity provided from a variety of sources including the CLEAVASE enzymes, the FEN-1 endonucleases (including RAD2 and XPG proteins), Taq DNA polymerase and *E. coli* DNA polymerase I.

The term "thermostable" when used in reference to an enzyme, such as a 5' nuclease, indicates that the enzyme is functional or active (*i.e.*, can perform catalysis) at an elevated temperature, *i.e.*, at about 55°C or higher.

The term "cleavage products" as used herein, refers to products generated by the reaction of a cleavage means with a cleavage structure (i.e., the treatment of a cleavage structure with a cleavage means).

5 The term "target nucleic acid" refers to a nucleic acid molecule containing a sequence that has at least partial complementarity with at least a probe oligonucleotide and may also have at least partial complementarity with an INVADER oligonucleotide. The target nucleic acid may comprise single- or double-stranded DNA or RNA, and may comprise nucleotide analogs, labels, and other modifications.

10 The term "probe oligonucleotide" refers to an oligonucleotide that interacts with a target nucleic acid to form a cleavage structure in the presence or absence of an INVADER oligonucleotide. When annealed to the target nucleic acid, the probe oligonucleotide and target form a cleavage structure and cleavage occurs within the probe oligonucleotide.

15 The term "non-target cleavage product" refers to a product of a cleavage reaction that is not derived from the target nucleic acid. As discussed above, in the methods of the present invention, cleavage of the cleavage structure generally occurs within the probe oligonucleotide. The fragments of the probe oligonucleotide generated by this target nucleic acid-dependent cleavage are "non-target cleavage products."

20 The term "INVADER oligonucleotide" refers to an oligonucleotide that hybridizes to a target nucleic acid at a location near the region of hybridization between a probe and the target nucleic acid, wherein the INVADER oligonucleotide comprises a portion (e.g., a chemical moiety, or nucleotide—whether complementary to that target or not) that overlaps with the region of hybridization between the probe and target. In some embodiments, the INVADER oligonucleotide contains sequences at its 3' end that are
25 substantially the same as sequences located at the 5' end of a probe oligonucleotide.

The term "substantially single-stranded" when used in reference to a nucleic acid substrate means that the substrate molecule exists primarily as a single strand of nucleic acid in contrast to a double-stranded substrate which exists as two strands of nucleic acid which are held together by inter-strand base pairing interactions.

30 The term "sequence variation" as used herein refers to differences in nucleic acid sequence between two nucleic acids. For example, a wild-type structural gene and a

mutant form of this wild-type structural gene may vary in sequence by the presence of single base substitutions and/or deletions or insertions of one or more nucleotides. These two forms of the structural gene are said to vary in sequence from one another. A second mutant form of the structural gene may exist. This second mutant form is said to vary in sequence from both the wild-type gene and the first mutant form of the gene.

The term "liberating" as used herein refers to the release of a nucleic acid fragment from a larger nucleic acid fragment, such as an oligonucleotide, by the action of, for example, a 5' nuclease such that the released fragment is no longer covalently attached to the remainder of the oligonucleotide.

The term " K_m " as used herein refers to the Michaelis-Menten constant for an enzyme and is defined as the concentration of the specific substrate at which a given enzyme yields one-half its maximum velocity in an enzyme catalyzed reaction.

The term "nucleotide" as used herein includes, but is not limited to, naturally occurring and/or synthetic nucleotides, nucleotide analogs, and nucleotide derivatives. For example, the term includes naturally occurring DNA or RNA monomers, nucleotides with backbone modifications such as peptide nucleic acid (PNA) (M. Egholm et al., Nature 365:566 [1993]), phosphorothioate DNA, phosphorodithioate DNA, phosphoramidate DNA, amide-linked DNA, MMI-linked DNA, 2'-O-methyl RNA, alpha-DNA and methylphosphonate DNA, nucleotides with sugar modifications such as 2'-O-methyl RNA, 2'-fluoro RNA, 2'-amino RNA, 2'-O-alkyl DNA, 2'-O-allyl DNA, 2'-O-alkynyl DNA, hexose DNA, pyranosyl RNA, and anhydrohexitol DNA, and nucleotides having base modifications such as C-5 substituted pyrimidines (substituents including fluoro-, bromo-, chloro-, iodo-, methyl-, ethyl-, vinyl-, formyl-, ethynyl-, propynyl-, alkynyl-, thiazoyl-, imidazolyl-, pyridyl-), 7-deazapurines with C-7 substituents including fluoro-, bromo-, chloro-, iodo-, methyl-, ethyl-, vinyl-, formyl-, alkynyl-, alkenyl-, thiazolyl-, imidazolyl-, pyridyl-), inosine and diaminopurine.

The term "base analog" as used herein refers to modified or non-naturally occurring bases such as 7-deaza purines (e.g., 7-deaza-adenine and 7-deaza-guanine); bases modified, for example, to provide altered interactions such as non-standard basepairing, including, but not limited to: IsoC, Iso G, and other modified bases and nucleotides described in U.S. Patent Nos. 5,432,272; 6,001,983; 6,037,120; 6,140,496;

5,912,340; 6,127,121 and 6,143,877, each of which is incorporated herein by reference in their entireties; heterocyclic base analogs based on the purine or pyrimidine ring systems, and other heterocyclic bases. Nucleotide analogs include base analogs and comprise modified forms of deoxyribonucleotides as well as ribonucleotides.

5 The term "polymorphic locus" is a locus present in a population that shows variation between members of the population (*e.g.*, the most common allele has a frequency of less than 0.95). In contrast, a "monomorphic locus" is a genetic locus at little or no variations seen between members of the population (generally taken to be a locus at which the most common allele exceeds a frequency of 0.95 in the gene pool of
10 the population).

 The term "microorganism" as used herein means an organism too small to be observed with the unaided eye and includes, but is not limited to bacteria, virus, protozoans, fungi, and ciliates.

 The term "microbial gene sequences" refers to gene sequences derived from a
15 microorganism.

 The term "bacteria" refers to any bacterial species including eubacterial and archaeobacterial species.

 The term "virus" refers to obligate, ultramicroscopic, intracellular parasites incapable of autonomous replication (*i.e.*, replication requires the use of the host cell's
20 machinery).

 The term "multi-drug resistant" or "multiple-drug resistant" refers to a microorganism that is resistant to more than one of the antibiotics or antimicrobial agents used in the treatment of said microorganism.

 The term "sample" in the present specification and claims is used in its broadest
25 sense. On the one hand it is meant to include a specimen or culture (*e.g.*, microbiological cultures). On the other hand, it is meant to include both biological and environmental samples. A sample may include a specimen of synthetic origin.

 Biological samples may be animal, including human, fluid, solid (*e.g.*, stool) or tissue, as well as liquid and solid food and feed products and ingredients such as dairy
30 items, vegetables, meat and meat by-products, and waste. Biological samples may be obtained from all of the various families of domestic animals, as well as feral or wild

animals, including, but not limited to, such animals as ungulates, bear, fish, lagamorphs, rodents, etc.

Environmental samples include environmental material such as surface matter, soil, water and industrial samples, as well as samples obtained from food and dairy processing instruments, apparatus, equipment, utensils, disposable and non-disposable items. These examples are not to be construed as limiting the sample types applicable to the present invention.

The term "source of target nucleic acid" refers to any sample that contains nucleic acids (RNA or DNA). Particularly preferred sources of target nucleic acids are biological samples including, but not limited to blood, saliva, cerebral spinal fluid, pleural fluid, milk, lymph, sputum and semen.

An oligonucleotide is said to be present in "excess" relative to another oligonucleotide (or target nucleic acid sequence) if that oligonucleotide is present at a higher molar concentration than the other oligonucleotide (or target nucleic acid sequence). When an oligonucleotide such as a probe oligonucleotide is present in a cleavage reaction in excess relative to the concentration of the complementary target nucleic acid sequence, the reaction may be used to indicate the amount of the target nucleic acid present. Typically, when present in excess, the probe oligonucleotide will be present at least a 100-fold molar excess; typically at least 1 pmole of each probe oligonucleotide would be used when the target nucleic acid sequence was present at about 10 fmoles or less.

A sample "suspected of containing" a first and a second target nucleic acid may contain either, both or neither target nucleic acid molecule.

The term "charge-balanced" oligonucleotide refers to an oligonucleotide (the input oligonucleotide in a reaction) that has been modified such that the modified oligonucleotide bears a charge, such that when the modified oligonucleotide is either cleaved (*i.e.*, shortened) or elongated, a resulting product bears a charge different from the input oligonucleotide (the "charge-unbalanced" oligonucleotide) thereby permitting separation of the input and reacted oligonucleotides on the basis of charge. The term "charge-balanced" does not imply that the modified or balanced oligonucleotide has a net neutral charge (although this can be the case). Charge-balancing refers to the design and

modification of an oligonucleotide such that a specific reaction product generated from this input oligonucleotide can be separated on the basis of charge from the input oligonucleotide.

For example, in an INVADER oligonucleotide-directed cleavage assay in which the probe oligonucleotide bears the sequence: 5' TTCTTTTCACCAGCGAGACGGG 3' (*i.e.*, SEQ ID NO:136 without the modified bases) and cleavage of the probe occurs between the second and third residues, one possible charge-balanced version of this oligonucleotide would be: 5' Cy3-AminoT-Amino-TCTTTTCACCAGCGAGAC GGG 3'. This modified oligonucleotide bears a net negative charge. After cleavage, the following oligonucleotides are generated: 5' Cy3-AminoT-Amino-T 3' and 5' CTTTTCACCAGCGAGACGGG 3' (residues 3-22 of SEQ ID NO:136). 5' Cy3-AminoT-Amino-T 3' bears a detectable moiety (the positively-charged Cy3 dye) and two amino-modified bases. The amino-modified bases and the Cy3 dye contribute positive charges in excess of the negative charges contributed by the phosphate groups and thus the 5' Cy3-AminoT-Amino-T 3' oligonucleotide has a net positive charge. The other, longer cleavage fragment, like the input probe, bears a net negative charge. Because the 5' Cy3-AminoT-Amino-T 3' fragment is separable on the basis of charge from the input probe (the charge-balanced oligonucleotide), it is referred to as a charge-unbalanced oligonucleotide. The longer cleavage product cannot be separated on the basis of charge from the input oligonucleotide as both oligonucleotides bear a net negative charge; thus, the longer cleavage product is not a charge-unbalanced oligonucleotide.

The term "net neutral charge" when used in reference to an oligonucleotide, including modified oligonucleotides, indicates that the sum of the charges present (*i.e.*, R-NH₃⁺ groups on thymidines, the N3 nitrogen of cytosine, presence or absence or phosphate groups, etc.) under the desired reaction or separation conditions is essentially zero. An oligonucleotide having a net neutral charge would not migrate in an electrical field.

The term "net positive charge" when used in reference to an oligonucleotide, including modified oligonucleotides, indicates that the sum of the charges present (*i.e.*, R-NH₃⁺ groups on thymidines, the N3 nitrogen of cytosine, presence or absence or phosphate groups, etc.) under the desired reaction conditions is +1 or greater. An

oligonucleotide having a net positive charge would migrate toward the negative electrode in an electrical field.

The term "net negative charge" when used in reference to an oligonucleotide, including modified oligonucleotides, indicates that the sum of the charges present (i.e., R-NH₃⁺ groups on thymidines, the N₃ nitrogen of cytosine, presence or absence or phosphate groups, etc.) under the desired reaction conditions is -1 or lower. An oligonucleotide having a net negative charge would migrate toward the positive electrode in an electrical field.

The term "polymerization means" or "polymerization agent" refers to any agent capable of facilitating the addition of nucleoside triphosphates to an oligonucleotide. Preferred polymerization means comprise DNA and RNA polymerases.

The term "ligation means" or "ligation agent" refers to any agent capable of facilitating the ligation (i.e., the formation of a phosphodiester bond between a 3'-OH and a 5' P located at the termini of two strands of nucleic acid). Preferred ligation means comprise DNA ligases and RNA ligases.

The term "reactant" is used herein in its broadest sense. The reactant can comprise, for example, an enzymatic reactant, a chemical reactant or light (e.g., ultraviolet light, particularly short wavelength ultraviolet light is known to break oligonucleotide chains). Any agent capable of reacting with an oligonucleotide to either shorten (i.e., cleave) or elongate the oligonucleotide is encompassed within the term "reactant."

The term "adduct" is used herein in its broadest sense to indicate any compound or element that can be added to an oligonucleotide. An adduct may be charged (positively or negatively) or may be charge-neutral. An adduct may be added to the oligonucleotide via covalent or non-covalent linkages. Examples of adducts include, but are not limited to, indodicarbocyanine dye amidites, amino-substituted nucleotides, ethidium bromide, ethidium homodimer, (1,3-propanediamino)propidium, (diethylenetriamino)propidium, thiazole orange, (N-N'-tetramethyl-1,3-propanediamino)propyl thiazole orange, (N-N'-tetramethyl-1,2-ethanediamino)propyl thiazole orange, thiazole orange-thiazole orange homodimer (TOTO), thiazole orange-thiazole blue heterodimer (TOTAB), thiazole orange-ethidium heterodimer 1 (TOED1),

thiazole orange-ethidium heterodimer 2 (TOED2) and fluorescein-ethidium heterodimer (FED), psoralens, biotin, streptavidin, avidin, etc.

Where a first oligonucleotide is complementary to a region of a target nucleic acid and a second oligonucleotide has complementary to the same region (or a portion of this region) a "region of sequence overlap" exists along the target nucleic acid. The degree of overlap will vary depending upon the nature of the complementarity (see, e.g., region "X" in Figs. 29 and 67 and the accompanying discussions).

As used herein, the term "purified" or "to purify" refers to the removal of contaminants from a sample. For example, recombinant CLEAVASE nucleases are expressed in bacterial host cells and the nucleases are purified by the removal of host cell proteins; the percent of these recombinant nucleases is thereby increased in the sample.

The term "recombinant DNA molecule" as used herein refers to a DNA molecule that comprises of segments of DNA joined together by means of molecular biological techniques.

The term "recombinant protein" or "recombinant polypeptide" as used herein refers to a protein molecule that is expressed from a recombinant DNA molecule.

As used herein the term "portion" when in reference to a protein (as in "a portion of a given protein") refers to fragments of that protein. The fragments may range in size from four amino acid residues to the entire amino acid sequence minus one amino acid (e.g., 4, 5, 6, . . . , n-1).

The term "nucleic acid sequence" as used herein refers to an oligonucleotide, nucleotide or polynucleotide, and fragments or portions thereof, and to DNA or RNA of genomic or synthetic origin that may be single or double stranded, and represent the sense or antisense strand. Similarly, "amino acid sequence" as used herein refers to peptide or protein sequence.

The term "peptide nucleic acid" ("PNA") as used herein refers to a molecule comprising bases or base analogs such as would be found in natural nucleic acid, but attached to a peptide backbone rather than the sugar-phosphate backbone typical of nucleic acids. The attachment of the bases to the peptide is such as to allow the bases to base pair with complementary bases of nucleic acid in a manner similar to that of an oligonucleotide. These small molecules, also designated anti gene agents, stop transcript

elongation by binding to their complementary strand of nucleic acid (Nielsen, *et al.* Anticancer Drug Des. 8:53 63 [1993]).

As used herein, the terms "purified" or "substantially purified" refer to molecules, either nucleic or amino acid sequences, that are removed from their natural environment, isolated or separated, and are at least 60% free, preferably 75% free, and most preferably 90% free from other components with which they are naturally associated. An "isolated polynucleotide" or "isolated oligonucleotide" is therefore a substantially purified polynucleotide.

As used herein, the term "fusion protein" refers to a chimeric protein containing the protein of interest (e.g., CLEAVASE BN/thrombin nuclease and portions or fragments thereof) joined to an exogenous protein fragment (the fusion partner which consists of a non CLEAVASE BN/thrombin nuclease protein). The fusion partner may enhance solubility of recombinant chimeric protein (e.g., the CLEAVASE BN/thrombin nuclease) as expressed in a host cell, may provide an affinity tag (e.g., a his-tag) to allow purification of the recombinant fusion protein from the host cell or culture supernatant, or both. If desired, the fusion protein may be removed from the protein of interest (e.g., CLEAVASE BN/thrombin nuclease or fragments thereof) by a variety of enzymatic or chemical means known to the art.

As used herein, the terms "chimeric protein" and "chimerical protein" refer to a single protein molecule that comprises amino acid sequences portions derived from two or more parent proteins. These parent molecules may be from similar proteins from genetically distinct origins, different proteins from a single organism, or different proteins from different organisms. By way of example but not by way of limitation, a chimeric structure-specific nuclease of the present invention may contain a mixture of amino acid sequences that have been derived from FEN-1 genes from two or more of the organisms having such genes, combined to form a non-naturally occurring nuclease. The term "chimerical" as used herein is not intended to convey any particular proportion of contribution from the naturally occurring genes, nor limit the manner in which the portions are combined. Any chimeric structure-specific nuclease constructs having cleavage activity as determined by the testing methods described herein are improved cleavage agents within the scope of the present invention.

The term "continuous strand of nucleic acid" as used herein is means a strand of nucleic acid that has a continuous, covalently linked, backbone structure, without nicks or other disruptions. The disposition of the base portion of each nucleotide, whether base-paired, single-stranded or mismatched, is not an element in the definition of a continuous strand. The backbone of the continuous strand is not limited to the ribose-phosphate or deoxyribose-phosphate compositions that are found in naturally occurring, unmodified nucleic acids. A nucleic acid of the present invention may comprise modifications in the structure of the backbone, including but not limited to phosphorothioate residues, phosphonate residues, 2' substituted ribose residues (*e.g.*, 2'-O-methyl ribose) and alternative sugar (*e.g.*, arabinose) containing residues.

The term "continuous duplex" as used herein refers to a region of double stranded nucleic acid in which there is no disruption in the progression of basepairs within the duplex (*i.e.*, the base pairs along the duplex are not distorted to accommodate a gap, bulge or mismatch with the confines of the region of continuous duplex). As used herein the term refers only to the arrangement of the basepairs within the duplex, without implication of continuity in the backbone portion of the nucleic acid strand. Duplex nucleic acids with uninterrupted basepairing, but with nicks in one or both strands are within the definition of a continuous duplex.

The term "duplex" refers to the state of nucleic acids in which the base portions of the nucleotides on one strand are bound through hydrogen bonding the their complementary bases arrayed on a second strand. The condition of being in a duplex form reflects on the state of the bases of a nucleic acid. By virtue of base pairing, the strands of nucleic acid also generally assume the tertiary structure of a double helix, having a major and a minor groove. The assumption of the helical form is implicit in the act of becoming duplexed.

The term "duplex dependent protein binding" refers to the binding of proteins to nucleic acid that is dependent on the nucleic acid being in a duplex, or helical form.

The term "duplex dependent protein binding sites or regions" as used herein refers to discrete regions or sequences within a nucleic acid that are bound with particular affinity by specific duplex-dependent nucleic acid binding proteins. This is in contrast to the generalized duplex-dependent binding of proteins that are not site-specific, such as

the histone proteins that bind chromatin with little reference to specific sequences or sites.

The term "protein-binding region" as used herein refers to a nucleic acid region identified by a sequence or structure as binding to a particular protein or class of proteins.

5 It is within the scope of this definition to include those regions that contain sufficient genetic information to allow identifications of the region by comparison to known sequences, but which might not have the requisite structure for actual binding (*e.g.*, a single strand of a duplex-dependent nucleic acid binding protein site). As used herein "protein binding region" excludes restriction endonuclease binding regions.

10 The term "complete double stranded protein binding region" as used herein refers to the minimum region of continuous duplex required to allow binding or other activity of a duplex-dependent protein. This definition is intended to encompass the observation that some duplex dependent nucleic acid binding proteins can interact with full activity with regions of duplex that may be shorter than a canonical protein binding region as
15 observed in one or the other of the two single strands. In other words, one or more nucleotides in the region may be allowed to remain unpaired without suppressing binding. As used here in, the term "complete double stranded binding region" refers to the minimum sequence that will accommodate the binding function. Because some such regions can tolerate non-duplex sequences in multiple places, although not necessarily
20 simultaneously, a single protein binding region might have several shorter sub-regions that, when duplexed, will be fully competent for protein binding.

The term "template" refers to a strand of nucleic acid on which a complementary copy is built from nucleoside triphosphates through the activity of a template-dependent nucleic acid polymerase. Within a duplex the template strand is, by convention, depicted
25 and described as the "bottom" strand. Similarly, the non-template strand is often depicted and described as the "top" strand.

The term "template-dependent RNA polymerase" refers to a nucleic acid polymerase that creates new RNA strands through the copying of a template strand as described above and which does not synthesize RNA in the absence of a template. This
30 is in contrast to the activity of the template-independent nucleic acid polymerases that

synthesize or extend nucleic acids without reference to a template, such as terminal deoxynucleotidyl transferase, or Poly A polymerase.

The term "ARRESTOR molecule" refers to an agent added to or included in an invasive cleavage reaction in order to stop one or more reaction components from participating in a subsequent action or reaction. This may be done by sequestering or inactivating some reaction component (*e.g.*, by binding or base-pairing a nucleic acid component, or by binding to a protein component). The term "ARRESTOR oligonucleotide" refers to an oligonucleotide included in an invasive cleavage reaction in order to stop or arrest one or more aspects of any reaction (*e.g.*, the first reaction and/or any subsequent reactions or actions; it is not intended that the ARRESTOR oligonucleotide be limited to any particular reaction or reaction step). This may be done by sequestering some reaction component (*e.g.*, base-pairing to another nucleic acid, or binding to a protein component). However, it is not intended that the term be so limited as to just situations in which a reaction component is sequestered.

As used herein, the term "kit" refers to any delivery system for delivering materials. In the context of reaction assays, such delivery systems include systems that allow for the storage, transport, or delivery of reaction reagents (*e.g.*, oligonucleotides, enzymes, etc. in the appropriate containers) and/or supporting materials (*e.g.*, buffers, written instructions for performing the assay etc.) from one location to another. For example, kits include one or more enclosures (*e.g.*, boxes) containing the relevant reaction reagents and/or supporting materials. As used herein, the term "fragmented kit" refers to a delivery systems comprising two or more separate containers that each contain a subportion of the total kit components. The containers may be delivered to the intended recipient together or separately. For example, a first container may contain an enzyme for use in an assay, while a second container contains oligonucleotides. The term "fragmented kit" is intended to encompass kits containing Analyte specific reagents (ASR's) regulated under section 520(e) of the Federal Food, Drug, and Cosmetic Act, but are not limited thereto. Indeed, any delivery system comprising two or more separate containers that each contains a subportion of the total kit components are included in the term "fragmented kit." In contrast, a "combined kit" refers to a delivery system containing all of the components of a reaction assay in a single container (*e.g.*, in a single

box housing each of the desired components). The term "kit" includes both fragmented and combined kits.

As used herein, the term "functional domain" refers to a region, or a part of a region, of a protein (e.g., an enzyme) that provides one or more functional characteristic of the protein. For example, a functional domain of an enzyme may provide, directly or indirectly, one or more activities of the enzyme including, but not limited to, substrate binding capability and catalytic activity. A functional domain may be characterized through mutation of one or more amino acids within the functional domain, wherein mutation of the amino acid(s) alters the associated functionality (as measured empirically in an assay) thereby indicating the presence of a functional domain.

As used herein, the term "heterologous functional domain" refers to a protein functional domain that is not in its natural environment. For example, a heterologous functional domain includes a functional domain from one enzyme introduced into another enzyme. A heterologous functional domain also includes a functional domain native to a protein that has been altered in some way (e.g., mutated, added in multiple copies, etc.). A heterologous functional domain may comprise a plurality of contiguous amino acids or may include two or more distal amino acids or amino acids fragments (e.g., two or more amino acids or fragments with intervening, non-heterologous, sequence). Heterologous functional domains are distinguished from endogenous functional domains in that the heterologous amino acid(s) are joined to or contain amino acid sequences that are not found naturally associated with the amino acid sequence in nature or are associated with a portion of a protein not found in nature.

As used herein, the term "altered functionality in a nucleic acid cleavage assay" refers to a characteristic of an enzyme that has been altered in some manner to differ from its natural state (e.g., to differ from how it is found in nature). Alterations include, but are not limited to, addition of a heterologous functional domain (e.g., through mutation or through creation of chimerical proteins). In some embodiments, the altered characteristic of the enzyme may be one that improves the performance of an enzyme in a nucleic acid cleavage assay. Types of improvement include, but are not limited to, improved nuclease activity (e.g., improved rate of reaction), improved substrate binding (e.g., increased or decreased binding of certain nucleic acid species [e.g., RNA or DNA] that produces a

desired outcome [e.g., greater specificity, improved substrate turnover, etc.]), and improved background specificity (e.g., less undesired product is produced). The present invention is not limited by the nucleic cleavage assay used to test improved functionality. However, in some preferred embodiments of the present invention, an invasive cleavage assay is used as the nucleic acid cleavage assay. In certain particularly preferred embodiments, an invasive cleavage assay utilizing an RNA target is used as the nucleic acid cleavage assay.

As used herein, the terms "N-terminal" and "C-terminal" in reference to polypeptide sequences refer to regions of polypeptides including portions of the N-terminal and C-terminal regions of the polypeptide, respectively. A sequence that includes a portion of the N-terminal region of polypeptide includes amino acids predominantly from the N-terminal half of the polypeptide chain, but is not limited to such sequences. For example, an N-terminal sequence may include an interior portion of the polypeptide sequence including bases from both the N-terminal and C-terminal halves of the polypeptide. The same applies to C-terminal regions. N-terminal and C-terminal regions may, but need not, include the amino acid defining the ultimate N-terminal and C-terminal ends of the polypeptide, respectively.

DESCRIPTION OF THE DRAWINGS

Figure 1 shows a schematic representation of sequential invasive cleavage reactions. In step A, an upstream INVADER oligonucleotide and a downstream probe combine with a target nucleic acid strand to form a cleavage structure. In step B, the portion of the cleaved signal probe from A combines with a second target nucleic acid strand and a labeled signal probe to form a second cleavage structure. In step C, cleavage of the labeled second cleavage structure yields a detectable signal.

Figure 2 shows schematic representations of several examples of invasive cleavage structures comprising RNA target strands (SEQ ID NO:141). Panel A depicts an INVADER oligonucleotide (SEQ ID NO:142) and probe (SEQ ID NO:143). Panel B depicts an INVADER oligonucleotide (SEQ ID NO:144) and probe (SEQ ID NO:143). Panel C depicts an INVADER oligonucleotide (SEQ ID NO:145) and probe (SEQ ID

NO:145). Panel D depicts an INVADER oligonucleotide (SEQ ID NO:145) and probe (SEQ ID NO:146).

Figure 3 shows schematic representations of two examples of structures that are not invasive cleavage structures labelled SEQ ID NOs:147-152.

5 Figure 4 shows a schematic representation of a configuration of invasive cleavage that is useful for detection of target sequence variations. In A, an invasive cleavage structure having overlap between the two probes is formed, and the arrow indicates that it is cleavable by the enzymes of the present invention. In B, variation of the target sequence removes a region of complementarity to the downstream probe and eliminates the overlap. The absence of an arrow in panel B indicates a reduced rate of cleavage of this structure compared to that diagrammed in panel A.

10 Figure 5 shows a diagram of the X-ray structure of a ternary complex of KlenTaqI with primer/template DNA in the polymerizing mode determined by Li *et al.* (Li *et al.*, Protein Sci., 7:1116 [1998]). Without intending to represent precise borders between features of the physical form, the portions referred to in the text as the "fingers", "thumb" and "palm" regions are loosely indicated by the circle, rectangle, and oval, respectively.

15 Figure 6 shows a schematic diagram of the DNA polymerase gene from *Thermus aquaticus*. Restriction sites used in these studies are indicated above. The approximate regions encoding various structural or functional domains of the protein are indicated by double-headed arrows, below.

20 Figure 7 shows a schematic diagram of the chimeric constructs comprising portions of the TaqPol gene and the TthPol gene. Open and shaded boxes denote TaqPol and TthPol sequences, respectively. The numbers correspond to the amino acid sequence of TaqPol. The 5' nuclease and polymerase domains of TaqPol and the palm, thumb, and fingers regions of the polymerase domain are indicated. The abbreviations for the restrictions sites used for recombination are as follows: E, EcoRI; N, NotI; Bs, BstBI; D, NdeI; B, BamHI; and S, Sall.

25 Figure 8A-H shows a comparison of the nucleotide structure of the polymerase genes isolated from *Thermus aquaticus* (SEQ ID NO:153), *Thermus flavus* (SEQ ID NO:154) and *Thermus thermophilus* (SEQ ID NO:155); the consensus sequence (SEQ ID NO:156) is shown at the top of each row.

Figure 9A-C shows a comparison of the amino acid sequence of the polymerase isolated from *Thermus aquaticus* (SEQ ID NO:157), *Thermus flavus* (SEQ ID NO:158), and *Thermus thermophilus* (SEQ ID NO:1); the consensus sequence (SEQ ID NO:159) is shown at the top of each row.

5 Figure 10 shows the sequences and proposed structures of substrates for the invasive signal amplification reaction with human IL-6 RNA target strand (SEQ ID NO:160) and upstream probe (SEQ ID NO:161). The cleavage site of the downstream probe (SEQ ID NO:162) is indicated by an arrow. Sequence of the IL-6 DNA target strand (SEQ ID NO:163) is shown below.

10 Figure 11 shows the image generated by a fluorescence imager showing the products of invasive cleavage assays using the indicated enzymes, and the IL-6 substrate of Figure 10 having either a DNA target strand (A) or an RNA target strand (B).

15 Figure 12 compares the cycling cleavage activities of Taq DN RX HT, Tth DN RX HT, and Taq-Tth chimerical enzymes with IL-6 substrate having an RNA target strand.

20 Figure 13 shows a comparison of the amino acid sequences of the BstI-BamHI fragments of TaqPol (SEQ ID NO:164) and TthPol (SEQ ID NO:165). Pairs of similar amino acids are shaded with light gray. Aligned amino acids that have a charge difference are shaded with dark gray. The numbers correspond to the amino acid sequence of TaqPol. Amino acids of TaqPol changed to the corresponding amino acids of TthPol by site-directed mutagenesis are indicated by (+).

25 Figure 14 compares the cycling cleavage activities of Taq DN RX HT, Taq-Tth chimerical enzymes, and chimerical enzymes having the indicated additional amino acid modifications, with IL-6 substrate having an RNA target strand.

 Figure 15 compares the cycling cleavage activities of Taq DN RX HT, Tth DN RX HT, and Taq DN RX HT having the indicated amino acid modifications, with IL-6 substrate having an RNA target strand.

 Figure 16 compares polymerization activities of TaqPol, TthPol, and Taq-Tth chimerical enzymes, and TaqPol having the indicated amino acid modifications.

Figure 17 shows a diagram of the X-ray structure of a ternary complex of Klentaq1 with primer/template DNA in the polymerizing mode determined by Li *et al.* (Li *et al.*, Protein Sci., 7:1116 [1998]). Amino acids G418 and E507 are indicated.

Figures 18 A-D show schematic diagrams of examples of substrates that may be used to measure various cleavage activities of enzymes. The substrates may be labeled, for example, with a fluorescent dye and a quenching moiety for FRET detection, as shown, to facilitate detection and measurement. The substrates of 18A and 18B are invasive cleavage structures having RNA and DNA target strands, respectively. 18C shows an example of an X-structure, and 18D shows an example of a hairpin structure, both of which may be used to assess the activity of enzymes on alternative structures that may be present in invasive cleavage reactions.

Figure 19 shows schematic diagrams of chimeric constructs comprising portions of the TaqPol gene and the TthPol gene. Open and shaded boxes denote TaqPol and TthPol sequences, respectively. The chimeras also include the DN, RX, and HT modifications. A table compares the cleavage activity of each protein on the indicated cleavage substrates.

Figure 20A shows a schematic diagram for an RNA containing invasive cleavage substrate. The 5' end of the target molecule (SEQ ID NO:166) is modified with biotin and blocked with streptavidin as described. The downstream probe (SEQ ID NO:167) with cleavage site is also shown. Panels B-D show analysis of the properties of the Taq DN RX HT G418K/E507Q mutant in cleavage of the shown substrate under conditions of varying reaction temperature, KCl concentration, and MgSO₄ concentration.

Figure 21 shows schematic diagrams for model substrates used to test enzymes for invasive cleavage activity. The molecule shown in 21A provides a DNA target strand (SEQ ID NO:168), while the model shown in 21B provides an RNA containing target strand (SEQ ID NO:167). Both 21A and B show downstream probe SEQ ID NO:166.

Figure 22 shows schematic diagrams for model substrates used to test enzymes for cleavage activity on alternative, non-invasive structures.

Figure 23 shows a schematic diagram for a model substrate used to test enzymes for invasive cleavage activity.

Figure 24 shows schematic diagrams for a model substrate used to test enzymes for invasive cleavage activity on RNA or DNA target strands.

Figure 25 compares the cycling cleavage activities of Tth DN RX HT, Taq 2M, TfiPol, Tsc Pol, and Tfi and Tsc-derived mutant enzymes.

Figure 26 depicts structures that may be employed to determine the ability of an enzyme to cleave a probe in the presence and the absence of an upstream oligonucleotide. Fig. 26 displays the sequence of oligonucleotide 89-15-1 (SEQ ID NO:212), oligonucleotide 81-69-5 (SEQ ID NO:213), oligonucleotide 81-69-4 (SEQ ID NO:214), oligonucleotide 81-69-3 (SEQ ID NO:215), oligonucleotide 81-69-2 (SEQ ID NO:216) and a portion of M13mp18 (SEQ ID NO:217).

Figure 27 shows the image generated by a fluorescence imager that shows the dependence of *Pfu* FEN-1 on the presence of an overlapping upstream oligonucleotide for specific cleavage of the probe.

Figure 28a shows the image generated by a fluorescence imager that compares the amount of product generated in a standard (*i.e.*, a non-sequential invasive cleavage reaction) and a sequential invasive cleavage reaction.

Figure 28b is a graph comparing the amount of product generated in a standard or basic (*i.e.*, a non-sequential invasive cleavage reaction) and a sequential invasive cleavage reaction ("INVADER sqrd") (y axis = fluorescence units; x axis = attomoles of target).

Figure 29 shows the image generated by a fluorescence imager that shows that the products of a completed sequential invasive cleavage reaction cannot cross contaminate a subsequent similar reaction.

Figure 30 shows the sequence of the oligonucleotide employed in an invasive cleavage reaction for the detection of HCMV viral DNA; Fig. 30 shows the sequence of oligonucleotide 89-76 (SEQ ID NO:218), oligonucleotide 89-44 (SEQ ID NO:219) and nucleotides 3057-3110 of the HCMV genome (SEQ ID NO:220).

Figure 31 shows the image generated by a fluorescence imager that shows the sensitive detection of HCMV viral DNA in samples containing human genomic DNA using an invasive cleavage reaction.

Figure 32 is a schematic that illustrates one embodiment of the present invention, where the cut probe from an initial invasive cleavage reaction is employed as the INVADERoligonucleotide in a second invasive cleavage reaction, and where an ARRESTOR oligonucleotide prevents participation of remaining uncut first probe in the cleavage of the second probe.

Figure 33 is a schematic that illustrates one embodiment of the present invention, where the cut probe from an initial invasive cleavage reaction is employed as an integrated INVADER-target complex in a second invasive cleavage reaction, and where an ARRESTOR oligonucleotide prevents participation of remaining uncut first probe in the cleavage of the second probe.

Figure 34 shows three images generated by a fluorescence imager showing that two different lengths of 2' O-methyl, 3' terminal amine-modified ARRESTOR oligonucleotide both reduce non-specific background cleavage of the secondary probe when included in the second step of a reaction where the cut probe from an initial invasive cleavage reaction is employed as an integrated INVADER-target complex in a second invasive cleavage reaction.

Figure 35A shows two images generated by a fluorescence imager showing the effects on nonspecific and specific cleavage signal of increasing concentrations of primary probe in the first step of a reaction where the cut probe from an initial invasive cleavage reaction is employed as the INVADER oligonucleotide in a second invasive cleavage reaction.

Figure 35B shows two images generated by a fluorescence imager showing the effects on nonspecific and specific cleavage signal of increasing concentrations of primary probe in the first step of a reaction, and inclusion of a 2' O-methyl, 3' terminal amine-modified ARRESTOR oligonucleotide in the second step of a reaction where the cut probe from an initial invasive cleavage reaction is employed as the INVADER oligonucleotide in a second invasive cleavage reaction.

Figure 35C shows shows a graph generated using the spreadsheet MICROSOFT EXCEL software, comparing the effects on nonspecific and specific cleavage signal of increasing concentrations of primary probe in the first step of a reaction, in the presence or absence of a 2' O-methyl, 3' terminal amine-modified ARRESTOR oligonucleotide in

the second step of a reaction where the cut probe from an initial invasive cleavage reaction is employed as the INVADER oligonucleotide in a second invasive cleavage reaction.

Figure 36A shows two images generated by a fluorescence imager showing the effects on nonspecific and specific cleavage signal of including an unmodified ARRESTOR oligonucleotide in the second step of a reaction where the cut probe from an initial invasive cleavage reaction is employed as the INVADER oligonucleotide in a second invasive cleavage reaction.

Figure 36B shows two images generated by a fluorescence imager showing the effects on nonspecific and specific cleavage signal of including a 3' terminal amine modified ARRESTOR oligonucleotide, a partially 2' O-methyl substituted, 3' terminal amine modified ARRESTOR oligonucleotide, or an entirely 2' O-methyl, 3' terminal amine modified ARRESTOR oligonucleotide in the second step of a reaction where the cut probe from an initial invasive cleavage reaction is employed as the INVADER oligonucleotide in a second invasive cleavage reaction.

Figure 37A shows two images generated by a fluorescence imager comparing the effects on nonspecific and specific cleavage signal of including ARRESTOR oligonucleotides of different lengths in the second step of a reaction where the cut probe from an initial invasive cleavage reaction is employed as the INVADER oligonucleotide in a second invasive cleavage reaction.

Figure 37B shows two images generated by a fluorescence imager comparing the effects on nonspecific and specific cleavage signal of including an arrestoer oligonucleotides of different lengths in the second step of a reaction where the cut probe from an initial invasive cleavage reaction is employed as the INVADER oligonucleotide in a second invasive cleavage reaction, and in which a longer variant of the secondary probe used in the reactions in Fig. 37A is tested.

Figure 37C shows a schematic diagram of a primary probe aligned with several ARRESTOR oligonucleotides of different lengths. The region of the primary probe that is complementary to the HBV target sequence is underlined. The ARRESTOR oligonucleotides are aligned with the probe by complementarity.

Figure 38 shows two images generated by a fluorescence imager comparing the effects on nonspecific and specific cleavage signal of including ARRESTOR oligonucleotides of different lengths in the second step of a reaction where the cut probe from an initial invasive cleavage reaction is employed as the INVADER oligonucleotide in a second invasive cleavage reaction, using secondary probes of two different lengths.

Figure 39 shows a graph of the calculated running average of a ten nucleotide stretch of the hUbiquitin RNA (the Ave(10) Index) derived from the SS-count output of an mfold analysis, expressed as a percentage of the total number of structures found by mfold that include a particular base, plotted against the position of the base.

Figure 40 shows an example microplate layout for an RNA INVADER assay comprising 40 samples, 6 standards, and a No Target Control.

Figure 41 shows INVADER assay components for use in detecting human (h), mouse (m), or rat (r) RNAs of the indicated genes or transcripts.

Figure 42 shows a computer display of an INVADERCREATOR Order Entry screen.

Figure 43 shows a computer display of an INVADERCREATOR Multiple SNP Design Selection screen.

Figure 44 shows a computer display of an INVADERCREATOR Designer Worksheet screen.

Figure 45 shows a computer display of an INVADERCREATOR Output Page screen.

Figure 46 shows a computer display of an INVADERCREATOR Printer Ready Output screen.

Figure 47 shows INVADER assay components (SEQ ID NOs:709-2640) for use in detecting RNA target nucleic acids. Components are grouped per RNA analyte to be detected. Where multiple probes, INVADER oligonucleotides, stacker oligonucleotides, ARRESTOR oligonucleotides, or other components are provided, any of the multiple components may be used, unless indicated otherwise. Unless indicated otherwise, oligonucleotides are presented 5'-3' orientation.

Figure 48 shows a chart showing the Ave(10) Index against base pair position.

DESCRIPTION OF THE INVENTION

Introduction

5 The present invention relates to methods and compositions for treating nucleic acid, and in particular, methods and compositions for detection and characterization of nucleic acid sequences and sequence changes.

10 In preferred embodiments, the present invention relates to means for cleaving a nucleic acid cleavage structure in a site-specific manner. While the present invention provides a variety of cleavage agents, in some embodiments, the present invention relates to a cleaving enzyme having 5' nuclease activity without interfering nucleic acid synthetic ability. In other embodiments, the present invention provides novel polymerases (*e.g.*, thermostable polymerases) possessing altered polymerase and/or nucleases activities.

15 For example, in some embodiments, the present invention provides 5' nucleases derived from thermostable DNA polymerases that exhibit altered DNA synthetic activity from that of native thermostable DNA polymerases. The 5' nuclease activity of the polymerase is retained while the synthetic activity is reduced or absent. Such 5' nucleases are capable of catalyzing the structure-specific cleavage of nucleic acids in the absence of interfering synthetic activity. The lack of synthetic activity during a cleavage reaction results in nucleic acid cleavage products of uniform size.

20 The novel properties of the nucleases of the invention form the basis of a method of detecting specific nucleic acid sequences. This method relies upon the amplification of the detection molecule rather than upon the amplification of the target sequence itself as do existing methods of detecting specific target sequences.

25 DNA polymerases (DNAPs), such as those isolated from *E. coli* or from thermophilic bacteria of the genus *Thermus* as well as other organisms, are enzymes that synthesize new DNA strands. Several of the known DNAPs contain associated nuclease activities in addition to the synthetic activity of the enzyme.

30 Some DNAPs are known to remove nucleotides from the 5' and 3' ends of DNA chains (Kornberg, DNA Replication, W.H. Freeman and Co., San Francisco, pp. 127-139 [1980]). These nuclease activities are usually referred to as 5' exonuclease and 3'

exonuclease activities, respectively. For example, the 5' exonuclease activity located in the N-terminal domain of several DNAPs participates in the removal of RNA primers during lagging strand synthesis during DNA replication and the removal of damaged nucleotides during repair. Some DNAPs, such as the *E. coli* DNA polymerase (DNAPEc1), also have a 3' exonuclease activity responsible for proof-reading during DNA synthesis (Kornberg, supra).

A DNAP isolated from *Thermus aquaticus*, termed Taq DNA polymerase (DNAPTaq), has a 5' exonuclease activity, but lacks a functional 3' exonucleolytic domain (Tindall and Kunkell, Biochem., 27:6008 [1988]). Derivatives of DNAPEc1 and DNAPTaq, respectively called the Klenow and Stoffel fragments, lack 5' exonuclease domains as a result of enzymatic or genetic manipulations (Brutlag et al., Biochem. Biophys. Res. Commun., 37:982 [1969]; Erlich et al., Science 252:1643 [1991]; Setlow and Kornberg, J. Biol. Chem., 247:232 [1972]).

The 5' exonuclease activity of DNAPTaq was reported to require concurrent synthesis (Gelfand, PCR Technology - Principles and Applications for DNA Amplification, H.A. Erlich, [Ed.], Stockton Press, New York, p. 19 [1989]). Although mononucleotides predominate among the digestion products of the 5' exonucleases of DNAPTaq and DNAPEc1, short oligonucleotides (≤ 12 nucleotides) can also be observed implying that these so-called 5' exonucleases can function endonucleolytically (Setlow, supra; Holland et al., Proc. Natl. Acad. Sci. USA 88:7276 [1991]).

In WO 92/06200, Gelfand et al. show that the preferred substrate of the 5' exonuclease activity of the thermostable DNA polymerases is displaced single-stranded DNA. Hydrolysis of the phosphodiester bond occurs between the displaced single-stranded DNA and the double-helical DNA with the preferred exonuclease cleavage site being a phosphodiester bond in the double helical region. Thus, the 5' exonuclease activity usually associated with DNAPs is a structure-dependent single-stranded endonuclease and is more properly referred to as a 5' nuclease. Exonucleases are enzymes that cleave nucleotide molecules from the ends of the nucleic acid molecule. Endonucleases, on the other hand, are enzymes that cleave the nucleic acid molecule at internal rather than terminal sites. The nuclease activity associated with some thermostable DNA polymerases cleaves endonucleolytically but this cleavage requires

contact with the 5' end of the molecule being cleaved. Therefore, these nucleases are referred to as 5' nucleases.

When a 5' nuclease activity is associated with a eubacterial Type A DNA polymerase, it is found in the one third N-terminal region of the protein as an independent functional domain. The C-terminal two-thirds of the molecule constitute the polymerization domain that is responsible for the synthesis of DNA. Some Type A DNA polymerases also have a 3' exonuclease activity associated with the two-third C-terminal region of the molecule.

The 5' exonuclease activity and the polymerization activity of DNAPs can be separated by proteolytic cleavage or genetic manipulation of the polymerase molecule. The Klenow or large proteolytic cleavage fragment of DNAPEc1 contains the polymerase and 3' exonuclease activity but lacks the 5' nuclease activity. The Stoffel fragment of DNAPTaq (DNAPStf) lacks the 5' nuclease activity due to a genetic manipulation that deleted the N-terminal 289 amino acids of the polymerase molecule (Erich et al., Science 252:1643 [1991]). WO 92/06200 describes a thermostable DNAP with an altered level of 5' to 3' exonuclease. U.S. Patent No. 5,108,892 describes a *Thermus aquaticus* DNAP without a 5' to 3' exonuclease. Thermostable DNA polymerases with lessened amounts of synthetic activity are available (Third Wave Technologies, Madison, WI) and are described in U.S. Pat. Nos. 5,541,311, 5,614,402, 5,795,763, 5,691,142, and 5,837,450, herein incorporated by reference in their entireties. The present invention provides 5' nucleases derived from thermostable Type A DNA polymerases that retain 5' nuclease activity but have reduced or absent synthetic activity. The ability to uncouple the synthetic activity of the enzyme from the 5' nuclease activity proves that the 5' nuclease activity does not require concurrent DNA synthesis as was previously reported (Gelfand, PCR Technology, supra).

In addition to the 5'-exonuclease domains of the DNA polymerase I proteins of Eubacteria, described above, 5' nucleases have been found associated with bacteriophage, eukaryotes and archaebacteria. Overall, all of the enzymes in this family display very similar substrate specificities, despite their limited level of sequence similarity. Consequently, enzymes suitable for use in the methods of the present invention may be isolated or derived from a wide array of sources.

A mammalian enzyme with functional similarity to the 5'-exonuclease domain of E. coli Pol I was isolated nearly 30 years ago (Lindahl, *et al.*, Proc Natl Acad Sci U S A 62(2): 597-603 [1969]). Later, additional members of this group of enzymes called flap endonucleases (FEN1) from Eukarya and Archaea were shown to possess a nearly identical structure specific activity (Harrington and Lieber. Embo J 13(5), 1235-46 [1994]; Murante *et al.*, J Biol Chem 269(2), 1191-6 [1994]; Robins, *et al.*, J Biol Chem 269(46), 28535-8 [1994]; Hosfield, *et al.*, J Biol Chem 273(42), 27154-61 [1998]), despite limited sequence similarity. The substrate specificities of the FEN1 enzymes, and the eubacterial and related bacteriophage enzymes have been examined and found to be similar for all enzymes (Lyamichev, *et al.*, Science 260(5109), 778-83 [1993], Harrington and Lieber, *supra*, Murante, *et al.*, *supra*, Hosfield, *et al.*, *supra*, Rao, *et al.*, J Bacteriol 180(20), 5406-12 [1998], Bhagwat, *et al.*, J. Biol Chem 272(45), 28523-30 [1997], Garforth and Sayers, Nucleic Acids Res 25(19), 3801-7 [1997]).

Using preformed substrates, many of the studies cited above determined that these nucleases leave a gap upon cleavage, leading the authors to speculate that DNA polymerase must then act to fill in that gap to generate a ligatable nick. A number of other 5' nucleases have been shown to leave a gap or overlap after cleavage of the same or similar flap substrates. It has since been determined that that all the structure-specific 5'-exonucleases leave a nick after cleavage if the substrate has an overlap between the upstream and downstream duplexes (Kaiser *et al.*, J. Biol .Chem. 274(30):21387-21394 [1999]). While duplexes having several bases of overlapping sequence can assume several different conformations through branch migration, it was determined that cleavage occurs in the conformation where the last nucleotide at the 3' end of the upstream strand is unpaired, with the cleavage rate being essentially the same whether the end of the upstream primer is A, C, G, or T. It was determined to be positional overlap between the 3' end of the upstream primer and downstream duplex, rather than sequence overlap, that provides optimal cleavage. In addition to allowing these enzymes to leave a nick after cleavage, the single base of overlap causes the enzymes to cleave several orders of magnitude faster than when a substrate lacks overlap (Kaiser *et al.*, *supra*).

Any of the 5' nucleases described below may find application in one or more embodiments of the methods described herein. 5' nucleases of particular utility in the

methods of present invention include but are not limited to polymerases from a *Thermus* species including, but not limited to, *Thermus aquaticus*, *Thermus flavus*, *Thermus thermophilus*, *Thermus filiformus*, and *Thermus scotoductus*, and altered polymerases.

Particularly preferred are altered polymerases exhibiting improved performance in detection assays based on the cleavage of a DNA member of an invasive cleavage structure that comprises an RNA target strand.

Chimerical polymerases may find application in one or more embodiments of the present invention, including but not limited to chimerical polymerases comprising one or more portions of one or more FEN nucleases including but are not limited to those of

Methanococcus jannaschii, *Methanobacterium thermoautotrophicum*, *Archaeoglobus veneficus*, *Sulfolobus solfataricus*, *Pyrobaculum aerophilum*, *Thermococcus litoralis*, *Archaeoglobus profundus*, *Acidianus brierlyi*, *Acidianus ambivalens*, *Desulfurococcus amylolyticus*, *Desulfurococcus mobilis*, *Pyrodictium brockii*, *Thermococcus gorgonarius*, *Thermococcus zilligii*, *Methanopyrus kandleri*, *Methanococcus igneus*, *Pyrococcus horikoshii*, and *Aeropyrum pernix*; particularly preferred FEN1 enzymes are chimerical *Archaeoglobus fulgidus* and *Pyrococcus furiosus*. Particularly preferred are altered polymerases exhibiting improved performance in detection assays based on the cleavage of a DNA member of an invasive cleavage structure that comprises an RNA target strand.

The detailed description of the invention is presented in the following sections:

- I. Detection of Specific Nucleic Acid Sequences Using 5' Nucleases in an INVADER Directed Cleavage Assay;
- II. Signal Enhancement By Incorporating The Products Of An Invasive Cleavage Reaction Into A Subsequent Invasive Cleavage Reaction;
- III. Effect of ARRESTOR Oligonucleotides on Signal and Background in Sequential Invasive Cleavage Reactions.
- IV. Improved Enzymes For Use In INVADER Oligonucleotide-Directed Cleavage Reactions Comprising RNA Targets;
- V. Reaction Design for INVADER Assay Detection of RNA Targets;
- VI. Kits for performing the RNA INVADER Assay; and

VII. The INVADER Assay for Direct Detection and Measurement of Specific RNA Analytes.

5 I. **Detection of Specific Nucleic Acid Sequences Using 5' Nucleases in an INVADER Directed Cleavage Assay**

1. INVADER Assay Reaction design

10 The present invention provides means for forming a nucleic acid cleavage structure that is dependent upon the presence of a target nucleic acid and cleaving the nucleic acid cleavage structure so as to release distinctive cleavage products. 5' nuclease activity, for example, is used to cleave the target-dependent cleavage structure and the resulting cleavage products are indicative of the presence of specific target nucleic acid sequences in the sample. When two strands of nucleic acid, or oligonucleotides, both
15 hybridize to a target nucleic acid strand such that they form an overlapping invasive cleavage structure, as described below, invasive cleavage can occur. Through the interaction of a cleavage agent (e.g., a 5' nuclease) and the upstream oligonucleotide, the cleavage agent can be made to cleave the downstream oligonucleotide at an internal site in such a way that a distinctive fragment is produced. Such embodiments have been
20 termed the INVADER assay (Third Wave Technologies) and are described in U.S. Patent Appl. Nos. 5,846,717, 5,985,557, 5,994,069, 6,001,567, and 6,090,543 and PCT Publications WO 97/27214 and WO 98/42873, herein incorporated by reference in their entireties.

25 The present invention further provides assays in which the target nucleic acid is reused or recycled during multiple rounds of hybridization with oligonucleotide probes and cleavage of the probes without the need to use temperature cycling (i.e., for periodic denaturation of target nucleic acid strands) or nucleic acid synthesis (i.e., for the polymerization-based displacement of target or probe nucleic acid strands). When a cleavage reaction is run under conditions in which the probes are continuously replaced
30 on the target strand (e.g. through probe-probe displacement or through an equilibrium between probe/target association and disassociation, or through a combination

comprising these mechanisms, [Reynaldo et al., J. Mol. Biol. 97: 511 (2000)]) multiple probes can hybridize to the same target, allowing multiple cleavages, and the generation of multiple cleavage products.

By the extent of its complementarity to a target nucleic acid strand, an oligonucleotide may be said to define a specific region of the target. In an invasive cleavage structure, the two oligonucleotides define and hybridize to regions of the target that are adjacent to one another (*i.e.*, regions without any additional region of the target between them). Either or both oligonucleotides may comprise additional portions that are not complementary to the target strand. In addition to hybridizing adjacently, in order to form an invasive cleavage structure, the 3' end of the upstream oligonucleotide must comprise an additional moiety. When both oligonucleotides are hybridized to a target strand to form a structure and such a 3' moiety is present on the upstream oligonucleotide within the structure, the oligonucleotides may be said to overlap, and the structure may be described as an overlapping, or invasive cleavage structure.

In one embodiment, the 3' moiety of the invasive cleavage structure is a single nucleotide. In this embodiment the 3' moiety may be any nucleotide (*i.e.*, it may be, but it need not be complementary to the target strand). In a preferred embodiment the 3' moiety is a single nucleotide that is not complementary to the target strand. In another embodiment, the 3' moiety is a nucleotide-like compound (*i.e.*, a moiety having chemical features similar to a nucleotide, such as a nucleotide analog or an organic ring compound; *See e.g.*, U.S. Pat. No. 5,985,557). In yet another embodiment the 3' moiety is one or more nucleotides that duplicate in sequence one or more nucleotides present at the 5' end of the hybridized region of the downstream oligonucleotide. In a further embodiment, the duplicated sequence of nucleotides of the 3' moiety is followed by a single nucleotide that is not further duplicative of the downstream oligonucleotide sequence, and that may be any other nucleotide. In yet another embodiment, the duplicated sequence of nucleotides of the 3' moiety is followed by a nucleotide-like compound, as described above.

The downstream oligonucleotide may have, but need not have, additional moieties attached to either end of the region that hybridizes to the target nucleic acid strand. In a preferred embodiment, the downstream oligonucleotide comprises a moiety at its 5' end

(*i.e.*, a 5' moiety). In a particularly preferred embodiment, said 5' moiety is a 5' flap or arm comprising a sequence of nucleotides that is not complementary to the target nucleic acid strand.

When an overlapping cleavage structure is formed, it can be recognized and
5 cleaved by a nuclease that is specific for this structure (*i.e.*, a nuclease that will cleave one or more of the nucleic acids in the overlapping structure based on recognition of this structure, rather than on recognition of a nucleotide sequence of any of the nucleic acids forming the structure). Such a nuclease may be termed a "structure-specific nuclease". In some embodiments, the structure-specific nuclease is a 5' nuclease. In a preferred
10 embodiment, the structure-specific nuclease is the 5' nuclease of a DNA polymerase. In another preferred embodiment, the DNA polymerase having the 5' nuclease is synthesis-deficient. In another preferred embodiment, the 5' nuclease is a FEN-1 endonuclease. In a particularly preferred embodiment, the 5' nuclease is thermostable.

In some embodiments, said structure-specific nuclease preferentially cleaves the
15 downstream oligonucleotide. In a preferred embodiment, the downstream oligonucleotide is cleaved one nucleotide into the 5' end of the region that is hybridized to the target within the overlapping structure. Cleavage of the overlapping structure at any location by a structure-specific nuclease produces one or more released portions or fragments of nucleic acid, termed "cleavage products."

20 In some embodiments, cleavage of an overlapping structure is performed under conditions wherein one or more of the nucleic acids in the structure can disassociate (*i.e.* un-hybridize, or melt) from the structure. In one embodiment, full or partial disassociation of a first cleavage structure allows the target nucleic acid to participate in the formation of one or more additional overlapping cleavage structures. In a preferred
25 embodiment, the first cleavage structure is partially disassociated. In a particularly preferred embodiment only the oligonucleotide that is cleaved disassociates from the first cleavage structure, such that it may be replaced by another copy of the same oligonucleotide. In some embodiments, said disassociation is induced by an increase in temperature, such that one or more oligonucleotides can no longer hybridize to the target
30 strand. In other embodiments, said disassociation occurs because cleavage of an oligonucleotide produces only cleavage products that cannot bind to the target strand

under the conditions of the reaction. In a preferred embodiment, conditions are selected wherein an oligonucleotide may associate with (*i.e.*, hybridize to) and disassociate from a target strand regardless of cleavage, and wherein the oligonucleotide may be cleaved when it is hybridized to the target as part of an overlapping cleavage structure. In a particularly preferred embodiment, conditions are selected such that the number of copies of the oligonucleotide that can be cleaved when part of an overlapping structure exceeds the number of copies of the target nucleic acid strand by a sufficient amount that when the first cleavage structure disassociates, the probability that the target strand will associate with an intact copy of the oligonucleotide is greater than the probability that it will associate with a cleaved copy of the oligonucleotide.

In some embodiments, cleavage is performed by a structure-specific nuclease that can recognize and cleave structures that do not have an overlap. In a preferred embodiment, cleavage is performed by a structure-specific nuclease having a lower rate of cleavage of nucleic acid structures that do not comprise an overlap, compared to the rate of cleavage of structures comprising an overlap. In a particularly preferred embodiment, cleavage is performed by a structure-specific nuclease having less than 1% of the rate of cleavage of nucleic acid structures that do not comprise an overlap, compared to the rate of cleavage of structures comprising an overlap.

In some embodiments it is desirable to detect the cleavage of the overlapping cleavage structure. Detection may be by analysis of cleavage products or by analysis of one or more of the remaining uncleaved nucleic acids. For convenience, the following discussion will refer to the analysis of cleavage products, but it will be appreciated by those skilled in the art that these methods may as easily be applied to analysis of the uncleaved nucleic acids in an invasive cleavage reaction. Any method known in the art for analysis of nucleic acids, nucleic acid fragments or oligonucleotides may be applied to the detection of cleavage products.

In one embodiment, the cleavage products may be identified by chemical content, *e.g.*, the relative amounts of each atom, each particular type of reactive group or each nucleotide base (Chargaff et al., *J. Biol. Chem.* 177: 405 [1949]) they contain. In this way, a cleavage product may be distinguished from a longer nucleic acid from which it was released by cleavage, or from other nucleic acids.

In another embodiment, the cleavage products may be distinguished by a particular physical attribute, including but not limited to length, mass, charge, or charge-to-mass ratio. In yet another embodiment, the cleavage product may be distinguished by a behavior that is related to a physical attribute, including but not limited to rate of rotation in solution, rate of migration during electrophoresis, coefficient of sedimentation in centrifugation, time of flight in MALDI-TOF mass spectrometry, migration rate or other behavior in chromatography, melting temperature from a complementary nucleic acid, or precipitability from solution.

Detection of the cleavage products may be through release of a label. Such labels may include, but are not limited to one or more of any of dyes, radiolabels such as ^{32}P or ^{35}S , binding moieties such as biotin, mass tags, such as metal ions or chemical groups, charge tags, such as polyamines or charged dyes, haptens such as digoxigenin, luminogenic, phosphorescent or fluorogenic moieties, and fluorescent dyes, either alone or in combination with moieties that can suppress or shift emission spectra, such as by fluorescence resonance energy transfer (FRET) or collisional fluorescence energy transfer.

In some embodiments, analysis of cleavage products may include physical resolution or separation, for example by electrophoresis, hybridization or by selective binding to a support, or by mass spectrometry methods such as MALDI-TOF. In other embodiments, the analysis may be performed without any physical resolution or separation, such as by detection of cleavage-induced changes in fluorescence as in FRET-based analysis, or by cleavage-induced changes in the rotation rate of a nucleic acid in solution as in fluorescence polarization analysis.

Cleavage products can be used subsequently in any reaction or read-out method that can make use of oligonucleotides. Such reactions include, but are not limited to, modification reactions, such as ligation, tailing with a template-independent nucleic acid polymerase and primer extension with a template-dependent nucleic acid polymerase. The modification of the cleavage products may be for purposes including, but not limited to, addition of one or more labels or binding moieties, alteration of mass, addition of specific sequences, or for any other purpose that would facilitate analysis of either the

cleavage products or analysis of any other by-product, result or consequence of the cleavage reaction.

Analysis of the cleavage products may involve subsequent steps or reactions that do not modify the cleavage products themselves. For example, cleavage products may be used to complete a functional structure, such as a competent promoter for *in vitro* transcription or another protein binding site. Analysis may include the step of using the completed structure for or to perform its function. One or more cleavage products may also be used to complete an overlapping cleavage structure, thereby enabling a subsequent cleavage reaction, the products of which may be detected or used by any of the methods described herein, including the participation in further cleavage reactions.

Certain preferred embodiments of the invasive cleavage reactions are provided in the following descriptions. In some embodiments, the methods of the present invention employ at least a pair of oligonucleotides that interact with a target nucleic acid to form a cleavage structure for a structure-specific nuclease. In some embodiments, the cleavage structure comprises i) a target nucleic acid that may be either single-stranded or double-stranded (when a double-stranded target nucleic acid is employed, it may be rendered single stranded, *e.g.*, by heating); ii) a first oligonucleotide, termed the "probe," that defines a first region of the target nucleic acid sequence by being the complement of that region; iii) a second oligonucleotide, termed the "INVADER oligonucleotide," the 5' part of which defines a second region of the same target nucleic acid sequence, adjacent to and downstream of the first target region, and the second part of which overlaps into the region defined by the first oligonucleotide.

It can be considered that the binding of these oligonucleotides in this embodiment divides the target nucleic acid into three distinct regions: one region that has complementarity to only the probe; one region that has complementarity only to the INVADER oligonucleotide; and one region that has complementarity to both oligonucleotides. As discussed above, in some preferred embodiments of the present invention, the overlap may comprise moieties other than overlapping complementary bases. Thus, in some embodiments, there is a physical, but not sequence, overlap between the INVADER and probe oligonucleotides, *i.e.*, in these latter embodiments,

there is not a region of the target nucleic acid that has complementarity to both oligonucleotides.

a) Oligonucleotide design

Design of these oligonucleotides (i.e., the INVADER oligonucleotide and the probe) is accomplished using practices that are standard in the art. For example, sequences that have self-complementarity, such that the resulting oligonucleotides would either fold upon themselves, or hybridize to each other at the expense of binding to the target nucleic acid, are generally avoided.

One consideration in choosing a length for these oligonucleotides is the complexity of the sample containing the target nucleic acid. For example, the human genome is approximately 3×10^9 basepairs in length. Any 10-nucleotide sequence will appear with a frequency of $1:4^{10}$, or 1:1,048,576 in a random string of nucleotides, which would be approximately 2,861 times in 3 billion basepairs. Clearly, an oligonucleotide of this length would have a poor chance of binding uniquely to a 10-nucleotide region within a target having a sequence the size of the human genome. If the target sequence were within a 3 kb plasmid, however, such an oligonucleotide might have a very reasonable chance of binding uniquely. By this same calculation it can be seen that an oligonucleotide of 16 nucleotides (i.e., a 16-mer) is the minimum length of a sequence that is mathematically likely to appear once in 3×10^9 basepairs. This level of specificity may also be provided by two or more shorter oligonucleotides if they are configured to bind in a cooperative fashion (i.e., such that they can produce the intended complex only if both or all are bound to their intended target sequences), wherein the combination of the short oligonucleotides provides the desired specificity. In one such embodiment, the cooperativity between the shorter oligonucleotides is by a coaxial stacking effect that can occur when the oligonucleotides hybridize to adjacent sites on a target nucleic acid. In another embodiment, the shorter oligonucleotides are connected to one another, either directly, or by one or more spacer regions. The short oligonucleotides thus connected may bind to distal regions of the target and may be used to bridge across regions of secondary structure in a target. Examples of such bridging oligonucleotides are described in PCT Publication WO 98/50403, herein incorporated by reference in its entirety.

A second consideration in choosing oligonucleotide length is the temperature range in which the oligonucleotides will be expected to function. A 16-mer of average base content (50% G-C bases) will have a calculated T_m of about 41°C, depending on, among other things, the concentration of the oligonucleotide and its target, the salt content of the reaction and the precise order of the nucleotides. As a practical matter, longer oligonucleotides are usually chosen to enhance the specificity of hybridization. Oligonucleotides 20 to 25 nucleotides in length are often used, as they are highly likely to be specific if used in reactions conducted at temperatures that are near their T_m s (within about 5 °C of the T_m). In addition, with calculated T_m s in the range of 50 to 70°C, such oligonucleotides (i.e., 20 to 25-mers) are appropriately used in reactions catalyzed by thermostable enzymes, which often display optimal activity near this temperature range.

The maximum length of the oligonucleotide chosen is also based on the desired specificity. One should avoid choosing sequences that are so long that they are either at a high risk of binding stably to partial complements, or that they cannot easily be dislodged when desired (e.g., failure to disassociate from the target once cleavage has occurred or failure to disassociate at a reaction temperature suitable for the enzymes and other materials in the reaction).

The first step of design and selection of the oligonucleotides for the INVADER oligonucleotide-directed cleavage is in accordance with these sample general principles. Considered as sequence-specific probes individually, each oligonucleotide may be selected according to the guidelines listed above. That is to say, each oligonucleotide will generally be long enough to be reasonably expected to hybridize only to the intended target sequence within a complex sample, usually in the 20 to 40 nucleotide range. Alternatively, because the INVADER oligonucleotide-directed cleavage assay depends upon the concerted action of these oligonucleotides, the composite length of the 2 oligonucleotides which span/bind to the target may be selected to fall within this range, with each of the individual oligonucleotides being in approximately the 13 to 17 nucleotide range. Such a design might be employed if a non-thermostable cleavage means were employed in the reaction, requiring the reactions to be conducted at a lower temperature than that used when thermostable cleavage means are employed. In some embodiments, it may be desirable to have these oligonucleotides bind multiple times

within a single target nucleic acid (*e.g.*, to bind to multiple variants or multiple similar sequences within a target). It is not intended that the method of the present invention be limited to any particular size of the probe or INVADER oligonucleotide.

The second step of designing an oligonucleotide pair for this assay is to choose the degree to which the upstream "INVADER" oligonucleotide sequence will overlap into the downstream "probe" oligonucleotide sequence, and consequently, the sizes into which the probe will be cleaved. In some preferred embodiments, the probe oligonucleotide can be made to "turn over," that is to say probe can be made to depart to allow the binding and cleavage of other copies of the probe molecule, without the requirements of thermal denaturation or displacement by polymerization. While in one embodiment of this assay probe turnover may be facilitated by an exonucleolytic digestion by the cleavage agent, in some preferred embodiments of the present invention turnover does not require this exonucleolytic activity. For example, in some embodiments, a reaction temperature and reaction conditions are selected so as to create an equilibrium wherein the probe hybridizes and disassociates from the target. In other embodiments, temperature and reaction conditions are selected so that unbound probe can initiate binding to the target strand and physically displace bound probe. In still other embodiments, temperature and reaction conditions are selected such that either or both mechanisms of probe replacement may occur in any proportion. The method of the present invention is not limited to any particular mechanism of probe replacement. By any mechanism, when the probe is bound to the target to form a cleavage structure, cleavage can occur. The continuous cycling of the probe on and off of the target allows multiple probes to bind and be cleaved for each copy of a target nucleic acid.

i) Non-sequence Overlaps

It has been determined that the relationship between the 3' end of the upstream oligonucleotide and the desired site of cleavage on the probe should be carefully designed. It is known that the preferred site of cleavage for the types of structure-specific endonucleases employed herein is one basepair into a duplex (Lyamichev *et al.*, *supra*). It was previously believed that the presence of an upstream oligonucleotide or primer allowed the cleavage site to be shifted away from this preferred site, into the single

stranded region of the 5' arm (Lyamichev *et al.*, *supra* and U.S. Patent No. 5,422,253). In contrast to this previously proposed mechanism, and while not limiting the present invention to any particular mechanism, it is believed that the nucleotide immediately 5', or upstream of the cleavage site on the probe (including miniprobe and mid-range probes) should be able to basepair with the target for efficient cleavage to occur. In the case of the present invention, this would be the nucleotide in the probe sequence immediately upstream of the intended cleavage site. In addition, as described herein, it has been observed that in order to direct cleavage to that same site in the probe, the upstream oligonucleotide should have its 3' base (*i.e.*, nt) immediately upstream of the intended cleavage site of the probe. In embodiments where the INVADER and probe oligonucleotides share a sequence overlap, this places the 3' terminal nucleotide of the upstream oligonucleotide and the base of the probe oligonucleotide 5' of the cleavage site in competition for pairing with the corresponding nucleotide of the target strand.

To examine the outcome of this competition (*i.e.* which base is paired during a successful cleavage event), substitutions were made in the probe and INVADER oligonucleotides such that either the probe or the INVADER oligonucleotide were mismatched with the target sequence at this position. The effects of both arrangements on the rates of cleavage were examined. When the INVADER oligonucleotide is unpaired at the 3' end, the rate of cleavage was not reduced. If this base was removed, however, the cleavage site was shifted upstream of the intended site. In contrast, if the probe oligonucleotide was not base-paired to the target just upstream of the site to which the INVADER oligonucleotide was directing cleavage, the rate of cleavage was dramatically reduced, suggesting that when a competition exists, the probe oligonucleotide was the molecule to be base-paired in this position.

It appears that the 3' end of the upstream INVADER oligonucleotide is unpaired during cleavage, and yet is important for accurate positioning of the cleavage. To examine which part(s) of the 3' terminal nucleotide are required for the positioning of cleavage, INVADER oligonucleotides were designed that terminated on this end with nucleotides that were altered in a variety of ways. Sugars examined included 2' deoxyribose with a 3' phosphate group, a dideoxyribose, 3' deoxyribose, 2' O-methyl ribose, arabinose and arabinose with a 3' phosphate. Abasic ribose, with and without 3'

phosphate were tested. Synthetic “universal” bases such as 3-nitropyrrole and 5-nitroindole on ribose sugars were tested. Finally, a base-like aromatic ring structure, acridine, linked to the 3' end of the previous nucleotide without a sugar group was tested. The results obtained support the conclusion that the aromatic ring of the base (at the 3' end of the INVADER oligonucleotide) is an important moiety for accomplishing the direction of cleavage to the desired site within the downstream probe. The 3' terminal moiety of the INVADER oligonucleotide need not be a base that is complementary to the target nucleic acid.

ii) Miniprobcs And Mid-Range Probes;

As discussed above, the INVADER oligonucleotide-directed cleavage assay may be performed using INVADER and probe oligonucleotides that have a length of about 13-25 nucleotides (typically 20-25 nucleotides). It is also contemplated that the oligonucleotides may themselves be composed of shorter oligonucleotide sequences that align along a target strand but that are not covalently linked. This is to say that there is a nick in the sugar-phosphate backbone of the composite oligonucleotide, but that there is no disruption in the progression of base-paired nucleotides in the resulting duplex. When short strands of nucleic acid align contiguously along a longer strand the hybridization of each is stabilized by the hybridization of the neighboring fragments because the basepairs can stack along the helix as though the backbone was in fact uninterrupted. This cooperativity of binding can give each segment a stability of interaction in excess of what would be expected for the segment hybridizing to the longer nucleic acid alone. One application of this observation has been to assemble primers for DNA sequencing, typically about 18 nucleotides long, from sets of three hexamer oligonucleotides that are designed to hybridize in this way (Kotler *et al.* Proc. Natl. Acad. Sci. USA 90:4241 [1993]). The resulting doubly-nicked primer can be extended enzymatically in reactions performed at temperatures that might be expected to disrupt the hybridization of hexamers, but not of 18-mers.

The use of composite or split oligonucleotides is applied with success in the INVADER-directed cleavage assay. For example, the probe oligonucleotide may be split into two oligonucleotides that anneal in a contiguous and adjacent manner along a target

oligonucleotide such that the downstream oligonucleotide (analogous to the probe) is assembled from two smaller pieces: a short segment of 6-10 nts (termed the "miniprobe"), that is to be cleaved in the course of the detection reaction, and an oligonucleotide that hybridizes immediately downstream of the miniprobe (termed the "stacker"), that serves to stabilize the hybridization of the probe. To form the cleavage structure, an upstream oligonucleotide (the INVADER oligonucleotide) is provided to direct the cleavage activity to the desired region of the miniprobe. Assembly of the probe from non-linked pieces of nucleic acid (*i.e.*, the miniprobe and the stacker) allows regions of sequences to be changed without requiring the re-synthesis of the entire proven sequence, thus improving the cost and flexibility of the detection system. In addition, the use of unlinked composite oligonucleotides makes the system more stringent in its requirement of perfectly matched hybridization to achieve signal generation, allowing this to be used as a sensitive means of detecting mutations or changes in the target nucleic acid sequences.

In one embodiment, the methods of the present invention employ at least three oligonucleotides that interact with a target nucleic acid to form a cleavage structure for a structure-specific nuclease. More specifically, the cleavage structure comprises i) a target nucleic acid that may be either single-stranded or double-stranded (when a double-stranded target nucleic acid is employed, it may be rendered single-stranded, *e.g.*, by heating); ii) a first oligonucleotide, termed the "stacker," that defines a first region of the target nucleic acid sequence by being the complement of that region.; iii) a second oligonucleotide, termed the "miniprobe," that defines a second region of the target nucleic acid sequence by being the complement of that region ; iv) a third oligonucleotide, termed the "INVADER oligonucleotide," the 5' part of which defines a third region of the same target nucleic acid sequence, adjacent to and downstream of the second target region, and the second or 3' part of which overlaps into the region defined by the second oligonucleotide As described above for embodiments that do not employ a stacker, the overlap region can represent a region where there is a physical, but not sequence, overlap between the INVADER and probe oligonucleotides.

In addition to the benefits cited above, the use of a composite design for the oligonucleotides that form the cleavage structure allows more latitude in the design of the reaction conditions for performing the INVADER-directed cleavage assay. When a longer probe (*e.g.*, 16-25 nt), as described above, is used for detection in reactions that are performed at temperatures below the T_m of that probe, the cleavage of the probe may play a significant role in destabilizing the duplex of which it is a part, thus allowing turnover and reuse of the recognition site on the target nucleic acid. In contrast, reaction temperatures that are at or above the T_m of the probe mean that the probe molecules are hybridizing and releasing from the target quite rapidly even without cleavage of the probe. When an upstream INVADER oligonucleotide and a cleavage agent are provided the probe will be specifically cleaved, but the cleavage will not be necessary to the turnover of the probe. When a long probe (*e.g.*, 16-25 nt) is used in this way the temperatures required to achieve this state is high, around 65 to 70°C for a 25-mer of average base composition. Requiring the use of such elevated temperatures limits the choice of cleavage agents to those that are very thermostable, and may contribute to background in the reactions, depending of the means of detection, through thermal degradation of the probe oligonucleotides. With miniprobcs, this latter mechanism of probe replacement may be accomplished at a lower temperature. Thus, shorter probes are preferred for embodiments using lower reaction temperatures.

The miniprobe of the present invention may vary in size depending on the desired application. In one embodiment, the probe may be relatively short compared to a standard probe (*e.g.*, 16-25 nt), in the range of 6 to 10 nucleotides. When such a short probe is used, reaction conditions can be chosen that prevent hybridization of the miniprobe in the absence of the stacker oligonucleotide. In this way a short probe can be made to assume the statistical specificity and selectivity of a longer sequence. In the event of a perturbation in the cooperative binding of the miniprobe and stacker nucleic acids, as might be caused by a mismatch within the short sequence that is otherwise complementary to the target nucleic acid or at the junction between the contiguous duplexes, this cooperativity can be lost, dramatically reducing the stability of the shorter duplex (*i.e.*, that of the miniprobe), and thus reducing the level of cleaved product in the assay of the present invention.

It is also contemplated that probes of intermediate size may be used. Such probes, in the 11 to 15 nucleotide range, may blend some of the features associated with the longer probes as originally described, these features including the ability to hybridize and be cleaved absent the help of a stacker oligonucleotide. At temperatures below the expected T_m of such probes, the mechanisms of turnover may be as discussed above for probes in the 20 nt range, and be dependent on the removal of the sequence in the overlap region for destabilization and cycling.

The mid-range probes may also be used at elevated temperatures, at or above their expected T_m , to allow melting rather than cleavage to promote probe turnover. In contrast to the longer probes described above, however, the temperatures required to allow the use of such a thermally driven turnover are much lower (about 40 to 60°C), thus preserving both the cleavage means and the nucleic acids in the reaction from thermal degradation. In this way, the mid-range probes may perform in some instances like the miniprobcs described above. In a further similarity to the miniprobcs, the accumulation of cleavage signal from a mid-range probe may be helped under some reaction conditions by the presence of a stacker.

To summarize, a standard long probe usually does not benefit from the presence of a stacker oligonucleotide downstream (the exception being cases where such an oligonucleotide may also disrupt structures in the target nucleic acid that interfere with the probe binding), and it may be used in conditions requiring several nucleotides to be removed to allow the oligonucleotide to release from the target efficiently. If temperature of the reaction is used to drive exchange of the probes, standard probes may require use of a temperature at which nucleic acids and enzymes are at higher risk of thermal degradation.

The miniprobe is very short and performs optimally in the presence of a downstream stacker oligonucleotide. The miniprobcs are well suited to reactions conditions that use the temperature of the reaction to drive rapid exchange of the probes on the target regardless of whether any bases have been cleaved. In reactions with sufficient amount of the cleavage means, the probes that do bind will be rapidly cleaved before they melt off.

The mid-range or midiprobe combines features of these probes and can be used in reactions like those favored by long probes, with longer regions of overlap to drive probe turnover at lower temperature. In a preferred embodiment, the midrange probes are used at temperatures sufficiently high that the probes are hybridizing to the target and releasing rapidly regardless of cleavage. The mid-range probe may have enhanced performance in the presence of a stacker under some circumstances.

The distinctions between the mini-, midi- (*i.e.*, mid-range) and long probes are not contemplated to be inflexible and based only on length. The performance of any given probe may vary with its specific sequence, the choice of solution conditions, the choice of temperature and the selected cleavage means.

The assemblage of oligonucleotides that comprises the cleavage structure of the present invention is sensitive to mismatches between the probe and the target. It is also contemplated that a mismatch between the INVADER oligonucleotide and the target may be used to distinguish related target sequences. In the 3-oligonucleotide system, comprising an INVADER, a probe and a stacker oligonucleotide, it is contemplated that mismatches may be located within any of the regions of duplex formed between these oligonucleotides and the target sequence. In a preferred embodiment, a mismatch to be detected is located in the probe. In a particularly preferred embodiment, the mismatch is in the probe, at the basepair immediately upstream (*i.e.*, 5') of the site that is cleaved when the probe is not mismatched to the target.

In another preferred embodiment, a mismatch to be detected is located within the region defined by the hybridization of a miniprobe. In a particularly preferred embodiment, the mismatch is in the miniprobe, at the basepair immediately upstream (*i.e.*, 5') of the site that is cleaved when the miniprobe is not mismatched to the target.

iii) Software for Oligonucleotide Design for the INVADER assay

The present invention provides systems and methods for the design of oligonucleotides for use in detection assays. In particular, the present invention provides systems and methods for the design of oligonucleotides that successfully hybridize to appropriate regions of target nucleic acids (e.g., regions of target nucleic acids that do not contain secondary structure) under the desired reaction conditions (e.g., temperature,

buffer conditions, etc.) for the detection assay. The systems and methods also allow for the design of multiple different oligonucleotides (e.g., oligonucleotides that hybridize to different portions of a target nucleic acid or that hybridize to two or more different target nucleic acids) that all function in the detection assay under the same or substantially the same reaction conditions. These systems and methods may also be used to design control samples that work under the experimental reaction conditions.

While the systems and methods of the present invention are not limited to any particular detection assay, the following description illustrates the invention when used in conjunction with the INVADER assay (Third Wave Technologies, Madison WI; See e.g., U.S. Pat. Nos. 5,846,717, 5,985,557, 5,994,069, and 6,001,567 and PCT Publications WO 97/27214 and WO 98/42873, incorporated herein by reference in their entireties) to detect a SNP. One skilled in the art will appreciate that specific and general features of this illustrative example are generally applicable to other detection assays, and for use in designing INVADER assays for purposes other than SNP detection (e.g., for DNA or RNA quantitation, for RNA splice junction detection, etc.). Further, it will be appreciated that all algorithms described herein can be applied as separate software elements, or calculations may be performed manually, for the design of any INVADER assay probe set without use of the INVADERCREATOR design system described below.

Oligonucleotide Design for the INVADER assay using the INVADERCREATOR program

In some embodiments where an oligonucleotide is designed for use in the INVADER assay to detect a SNP, the sequence(s) of interest are entered into the INVADERCREATOR program (Third Wave Technologies, Madison, WI). As described above, sequences may be input for analysis from any number of sources, either directly into the computer hosting the INVADERCREATOR program, or via a remote computer linked through a communication network (e.g., a LAN, Intranet or Internet network).

The program designs probes for both the sense and antisense strand. Strand selection is generally based upon the ease of synthesis, minimization of secondary structure formation, and manufacturability. In some embodiments, the user chooses the strand for sequences to be designed for. In other embodiments, the software automatically selects the strand. By incorporating thermodynamic parameters for optimum probe cycling and

signal generation (Allawi and SantaLucia, Biochemistry, 36:10581 [1997]), oligonucleotide probes may be designed to operate at a pre-selected assay temperature (e.g., 63°C). Based on these criteria, a final probe set (e.g., primary probes for 2 alleles and an INVADER oligonucleotide) is selected.

5 In some embodiments, the INVADERCREATOR system is a web-based program with secure site access that contains a link to BLAST (available at the National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health web site) and that can be linked to RNAstructure (Mathews *et al.*, RNA 5:1458 [1999]), a software program that incorporates mfold (Zuker, Science, 244:48 [1989]).

10 RNAstructure tests the proposed oligonucleotide designs generated by INVADERCREATOR for potential uni- and bimolecular complex formation. INVADERCREATOR is open database connectivity (ODBC)-compliant and uses the Oracle database for export/integration. The INVADERCREATOR system was configured with Oracle to work well with UNIX systems, as most genome centers are
15 UNIX-based.

 In some embodiments, the INVADERCREATOR analysis is provided on a separate server (e.g., a Sun server) so it can handle analysis of large batch jobs. For example, a customer can submit up to 2,000 SNP sequences in one email. The server passes the batch of sequences on to the INVADERCREATOR software, and, when
20 initiated, the program designs SNP sets. In some embodiments, probe set designs are returned to the user within 24 hours of receipt of the sequences.

 In some preferred embodiments, each INVADER assay reaction includes at least two target sequence-specific, unlabeled oligonucleotides for the primary reaction: an upstream INVADER oligonucleotide and a downstream Probe oligonucleotide. The
25 INVADER oligonucleotide is generally designed to bind stably at the reaction temperature, while the probe is designed to freely associate and disassociate with the target strand, with cleavage occurring only when an uncut probe hybridizes adjacent to an overlapping INVADER oligonucleotide. In some embodiments, the probe includes a 5' flap or "arm" that is not complementary to the target, and this flap is released from the
30 probe when cleavage occurs. In some embodiments, the released flap participates as an INVADER oligonucleotide in a secondary reaction.

The following discussion provides one example of how a user interface for an INVADERCREATOR program may be configured.

The user opens a work screen (Figure 42), *e.g.*, by clicking on an icon on a desktop display of a computer (*e.g.*, a Windows desktop). The user enters information related to the target sequence for which an assay is to be designed. In some embodiments, the user enters a target sequence. In other embodiments, the user enters a code or number that causes retrieval of a sequence from a database. In still other embodiments, additional information may be provided, such as the user's name, an identifying number associated with a target sequence, and/or an order number. In preferred embodiments, the user indicates (*e.g.* via a check box or drop down menu) that the target nucleic acid is DNA or RNA. In other preferred embodiments, the user indicates the species from which the nucleic acid is derived. In particularly preferred embodiments, the user indicates whether the design is for monoplex (*i.e.*, one target sequence or allele per reaction) or multiplex (*i.e.*, multiple target sequences or alleles per reaction) detection. When the requisite choices and entries are complete, the user starts the analysis process. In one embodiment, the user clicks a "Go Design It" button to continue.

In some embodiments, the software validates the field entries before proceeding. In some embodiments, the software verifies that any required fields are completed with the appropriate type of information. In other embodiments, the software verifies that the input sequence meets selected requirements (*e.g.*, minimum or maximum length, DNA or RNA content). If entries in any field are not found to be valid, an error message or dialog box may appear. In preferred embodiments, the error message indicates which field is incomplete and/or incorrect. Once a sequence entry is verified, the software proceeds with the assay design.

In some embodiments, the information supplied in the order entry fields specifies what type of design will be created. In preferred embodiments, the target sequence and multiplex check box specify which type of design to create. Design options include but are not limited to SNP assay, Multiplexed SNP assay (*e.g.*, wherein probe sets for different alleles are to be combined in a single reaction), Multiple SNP assay (*e.g.*,

wherein an input sequence has multiple sites of variation for which probe sets are to be designed), and Multiple Probe Arm assays.

In some embodiments, the INVADERCREATOR software is started via a Web Order Entry (WebOE) process (*i.e.*, through an Intra/Internet browser interface) and these parameters are transferred from the WebOE via applet <param> tags, rather than entered through menus or check boxes.

In the case of Multiple SNP Designs, the user chooses two or more designs to work with. In some embodiments, this selection opens a new screen view (*e.g.*, a Multiple SNP Design Selection view Figure 43). In some embodiments, the software creates designs for each locus in the target sequence, scoring each, and presents them to the user in this screen view. The user can then choose any two designs to work with. In some embodiments, the user chooses a first and second design (*e.g.*, via a menu or buttons) and clicks a "Go Design It" button to continue.

To select a probe sequence that will perform optimally at a pre-selected reaction temperature, the melting temperature (T_m) of the SNP to be detected is calculated using the nearest-neighbor model and published parameters for DNA duplex formation (Allawi and SantaLucia, Biochemistry, 36:10581 [1997]). In embodiments wherein the target strand is RNA, parameters appropriate for RNA/DNA heteroduplex formation may be used. Because the assay's salt concentrations are often different than the solution conditions in which the nearest-neighbor parameters were obtained (1M NaCl and no divalent metals), and because the presence and concentration of the enzyme influence optimal reaction temperature, an adjustment should be made to the calculated T_m to determine the optimal temperature at which to perform a reaction. One way of compensating for these factors is to vary the value provided for the salt concentration within the melting temperature calculations. This adjustment is termed a 'salt correction'. As used herein, the term "salt correction" refers to a variation made in the value provided for a salt concentration for the purpose of reflecting the effect on a T_m calculation for a nucleic acid duplex of a non-salt parameter or condition affecting said duplex. Variation of the values provided for the strand concentrations will also affect the outcome of these calculations. By using a value of 0.5 M NaCl (SantaLucia, Proc Natl Acad Sci U S A, 95:1460 [1998]) and strand concentrations of about 1 mM of the probe and 1 fM target,

the algorithm used for calculating probe-target melting temperature has been adapted for use in predicting optimal INVADER assay reaction temperature. For a set of 30 probes, the average deviation between optimal assay temperatures calculated by this method and those experimentally determined is about 1.5 °C.

5 The length of the downstream probe analyte-specific region (ASR) is defined by the temperature selected for running the reaction (*e.g.*, 63°C). Starting from the position of the variant nucleotide on the target DNA (the target base that is paired to the probe nucleotide 5' of the intended cleavage site), and adding on the 3' end, an iterative procedure is used by which the length of the target-binding region of the probe is
10 increased by one base pair at a time until a calculated optimal reaction temperature (T_m plus salt correction to compensate for enzyme effect) matching the desired reaction temperature is reached. The non-complementary arm of the probe is preferably selected to allow the secondary reaction to cycle at the same reaction temperature. The entire probe oligonucleotide is screened using programs such as *mfold* (Zuker, Science, 244: 48
15 [1989]) or Oligo 5.0 (Rychlik and Rhoads, Nucleic Acids Res, 17: 8543 [1989]) for the possible formation of dimer complexes or secondary structures that could interfere with the reaction. The same principles are also followed for INVADER oligonucleotide design. Briefly, starting from the position N on the target DNA, the 3' end of the INVADER oligonucleotide is designed to have a nucleotide not complementary to either
20 allele suspected of being contained in the sample to be tested. The mismatch does not adversely affect cleavage (Lyamichev *et al.*, Nature Biotechnology, 17: 292 [1999]), and it can enhance probe cycling, presumably by minimizing coaxial stabilization effects between the two probes. Additional residues complementary to the target DNA starting from residue N-1 are then added in the 5' direction until the stability of the
25 INVADERoligonucleotide-target hybrid exceeds that of the probe (and therefore the planned assay reaction temperature), generally by 15-20 °C.

 In some embodiments, the released cleavage fragment from a primary reaction is to be used in a secondary reaction. It is one aspect of the assay design that the all of the probe sequences may be selected to allow the primary and secondary reactions to occur at
30 the same optimal temperature, so that the reaction steps can run simultaneously. In an alternative embodiment, the probes may be designed to operate at different optimal

temperatures, so that the reaction steps are not simultaneously at their temperature optima.

In some embodiments, the software provides the user an opportunity to change various aspects of the design including but not limited to: probe, target and INVADER oligonucleotide temperature optima and concentrations; blocking groups; probe arms; dyes, capping groups and other adducts; individual bases of the probes and targets (*e.g.*, adding or deleting bases from the end of targets and/or probes, or changing internal bases in the INVADER and/or probe and/or target oligonucleotides). In some embodiments, changes are made by selection from a menu. In other embodiments, changes are entered into text or dialog boxes. In preferred embodiments, this option opens a new screen (*e.g.*, a Designer Worksheet view, Figure 44).

In some embodiments, the software provides a scoring system to indicate the quality (*e.g.*, the likelihood of performance) of the assay designs. In one embodiment, the scoring system includes a starting score of points (*e.g.*, 100 points) wherein the starting score is indicative of an ideal design, and wherein design features known or suspected to have an adverse affect on assay performance are assigned penalty values. Penalty values may vary depending on assay parameters other than the sequences, including but not limited to the type of assay for which the design is intended (*e.g.*, monoplex, multiplex) and the temperature at which the assay reaction will be performed. The following example provides illustrative scoring criteria for use with some embodiments of the INVADER assay based on an intelligence defined by experimentation. Examples of design features that may incur score penalties include but are not limited to the following [penalty values are indicated in brackets, first number is for lower temperature assays (*e.g.*, 62-64 °C), second is for higher temperature assays (*e.g.*, 65-66 °C)]:

1. [100:100] 3' end of INVADER oligonucleotide resembles the probe arm:

ARM SEQUENCE:

PENALTY AWARDED IF INVADER

OLIGONUCLEOTIDE ENDS IN:

Arm 1: CGCGCCGAGG

5'...GAGGX or 5'...GAGGXX

Arm 2: ATGACGTGGCAGAC

5'...CAGACX or 5'...CAGACXX

Arm 3: ACGGACGCGGAG 5'...GGAGX or 5'...GGAGXX

Arm 4: TCCGCGCGTCC 5'...GTCCX or 5'...GTCCXX

2. [70:70] a probe has 5-base stretch (*i.e.*, 5 of the same base in a row) containing the polymorphism;

5 3. [60:60] a probe has 5-base stretch adjacent to the polymorphism;

4. [50:50] a probe has 5-base stretch one base from the polymorphism;

5. [40:40] a probe has 5-base stretch two bases from the polymorphism;

6. [50:50] probe 5-base stretch is of Gs – additional penalty;

7. [100:100] a probe has 6-base stretch anywhere;

10 8. [90:90] a two or three base sequence repeats at least four times;

9. [100:100] a degenerate base occurs in a probe;

10. [60:90] probe hybridizing region is short (13 bases or less for designs 65-67°C; 12 bases or less for designs 62-64°C)

11. [40:90] probe hybridizing region is long (29 bases or more for designs 65-67°C, 28

15 bases or more for designs 62-64°C)

12. [5:5] probe hybridizing region length – per base additional penalty

13. [80:80] Ins/Del design with poor discrimination in first 3 bases after probe arm

14. [100:100] calculated INVADER oligonucleotide T_m within 7.5°C of probe target T_m (designs 65-67°C with INVADER oligonucleotide less than ≤ 70.5°C, designs 62-

20 64°C with INVADER oligonucleotide ≤ 69.5°C

15. [20:20] calculated probes T_ms differ by more than 2.0°C

16. [100:100] a probe has calculated T_m 2°C less than its target T_m

17. [10:10] target of one strand 8 bases longer than that of other strand

18. [30:30] INVADER oligonucleotide has 6-base stretch anywhere – initial penalty

25 19. [70:70] INVADER oligonucleotide 6-base stretch is of Gs - additional penalty

20. [15:15] probe hybridizing region is 14, 15 or 24-28 bases long (65-67°C) or 13,14 or 26,27 bases long (62-64°C)

21. [15:15] a probe has a 4-base stretch of Gs containing the polymorphism

30 In particularly preferred embodiments, temperatures for each of the oligonucleotides in the designs are recomputed and scores are recomputed as changes are

made. In some embodiments, score descriptions can be seen by clicking a "descriptions" button. In some embodiments, a BLAST search option is provided. In preferred embodiments, a BLAST search is done by clicking a "BLAST Design" button. In some embodiments, this action brings up a dialog box describing the BLAST process. In preferred embodiments, the BLAST search results are displayed as a highlighted design on a Designer Worksheet.

In some embodiments, a user accepts a design by clicking an "Accept" button. In other embodiments, the program approves a design without user intervention. In preferred embodiments, the program sends the approved design to a next process step (e.g., into production; into a file or database). In some embodiments, the program provides a screen view (e.g., an Output Page, Figure 45), allowing review of the final designs created and allowing notes to be attached to the design. In preferred embodiments, the user can return to the Designer Worksheet (e.g., by clicking a "Go Back" button) or can save the design (e.g., by clicking a "Save It" button) and continue (e.g., to submit the designed oligonucleotides for production).

In some embodiments, the program provides an option to create a screen view of a design optimized for printing (e.g., a text-only view) or other export (e.g., an Output view, Figure 46). In preferred embodiments, the Output view provides a description of the design particularly suitable for printing, or for exporting into another application (e.g., by copying and pasting into another application). In particularly preferred embodiments, the Output view opens in a separate window.

The present invention is not limited to the use of the INVADERCREATOR software. Indeed, a variety of software programs are contemplated and are commercially available, including, but not limited to GCG Wisconsin Package (Genetics computer Group, Madison, WI) and Vector NTI (Informax, Rockville, Maryland).

b) Design Of The Reaction Conditions

Target nucleic acids (e.g., RNA and DNA) that may be analyzed using the methods of the present invention that employ a 5' nuclease or other appropriate cleavage agents. Such nucleic acids may be obtained using standard molecular biological techniques. For example, nucleic acids (RNA or DNA) may be isolated from a tissue

sample (e.g., a biopsy specimen), tissue culture cells, samples containing bacteria and/or viruses (including cultures of bacteria and/or viruses), etc. The target nucleic acid may also be transcribed *in vitro* from a DNA template or may be chemically synthesized or amplified in by polymerase chain reaction. Furthermore, nucleic acids may be isolated from an organism, either as genomic material or as a plasmid or similar extrachromosomal DNA, or they may be a fragment of such material generated by treatment with a restriction endonuclease or other cleavage agent, or a shearing force, or it may be synthetic.

Assembly of the target, probe, and INVADER oligonucleotide nucleic acids into the cleavage reaction of the present invention uses principles commonly used in the design of oligonucleotide-based enzymatic assays, such as dideoxynucleotide sequencing and polymerase chain reaction (PCR). As is done in these assays, the oligonucleotides are provided in sufficient excess that the rate of hybridization to the target nucleic acid is very rapid. These assays are commonly performed with 50 fmoles to 2 pmoles of each oligonucleotide per microliter of reaction mixture, although they are not necessarily limited to this range. In the Examples described herein, amounts of oligonucleotides ranging from 250 fmoles to 5 pmoles per microliter of reaction volume were used. These values were chosen for the purpose of ease in demonstration and are not intended to limit the performance of the present invention to these concentrations. Other (e.g., lower) oligonucleotide concentrations commonly used in other molecular biological reactions are also contemplated.

It is desirable that an INVADER oligonucleotide be immediately available to direct the cleavage of each probe oligonucleotide that hybridizes to a target nucleic acid. In some embodiments described herein, the INVADER oligonucleotide is provided in excess over the probe oligonucleotide. While this is an effective means of making the INVADER oligonucleotide immediately available in such embodiments it is not intended that the practice of the present invention be limited to conditions wherein the INVADER oligonucleotide is in excess over the probe, or to any particular ratio of INVADER-to-probe (e.g., in some preferred embodiments described herein, the probe is provided in excess over the INVADER oligonucleotide). Another means of assuring the presence of an INVADER oligonucleotide whenever a probe binds to a target nucleic acid is to

design the INVADER oligonucleotide to hybridize more stably to the target, *i.e.*, to have a higher T_m than the probe. This can be accomplished by any of the means of increasing nucleic acid duplex stability discussed herein (*e.g.*, by increasing the amount of complementarity to the target nucleic acid).

5 Buffer conditions should be chosen that are compatible with both the oligonucleotide/target hybridization and with the activity of the cleavage agent. The optimal buffer conditions for nucleic acid modification enzymes, and particularly DNA modification enzymes, generally included enough mono- and di-valent salts to allow association of nucleic acid strands by base-pairing. If the method of the present invention is performed using an enzymatic cleavage agent other than those specifically described here, the reactions may generally be performed in any such buffer reported to be optimal for the nuclease function of the cleavage agent. In general, to test the utility of any cleavage agent in this method, test reactions are performed wherein the cleavage agent of interest is tested in the MOPS/MnCl₂/KCl buffer or Mg-containing buffers described 10 herein and in whatever buffer has been reported to be suitable for use with that agent, in a manufacturer's data sheet, a journal article, or in personal communication.

The products of the INVADER oligonucleotide-directed cleavage reaction are fragments generated by structure-specific cleavage of the input oligonucleotides. The resulting cleaved and/or uncleaved oligonucleotides may be analyzed and resolved by a number of methods including, but not limited to, electrophoresis (on a variety of supports including acrylamide or agarose gels, paper, etc.), chromatography, fluorescence polarization, mass spectrometry and chip hybridization. In some Examples the invention is illustrated using electrophoretic separation for the analysis of the products of the cleavage reactions. However, it is noted that the resolution of the cleavage products is 20 not limited to electrophoresis. Electrophoresis is chosen to illustrate the method of the invention because electrophoresis is widely practiced in the art and is easily accessible to the average practitioner. In other Examples, the invention is illustrated without electrophoresis or any other resolution of the cleavage products.

The probe and INVADER oligonucleotides may contain a label to aid in their 30 detection following the cleavage reaction. The label may be a radioisotope (*e.g.*, a ³²P or ³⁵S-labelled nucleotide) placed at either the 5' or 3' end of the oligonucleotide or

alternatively, the label may be distributed throughout the oligonucleotide (i.e., a uniformly labeled oligonucleotide). The label may be a nonisotopic detectable moiety, such as a fluorophore, that can be detected directly, or a reactive group that permits specific recognition by a secondary agent. For example, biotinylated oligonucleotides

5 may be detected by probing with a streptavidin molecule that is coupled to an indicator (e.g., alkaline phosphatase or a fluorophore) or a hapten such as dioxigenin may be detected using a specific antibody coupled to a similar indicator. The reactive group may also be a specific configuration or sequence of nucleotides that can bind or otherwise interact with a secondary agent, such as another nucleic acid, and enzyme, or an antibody.

10 In some embodiments, a probe is labeled with fluorescing moiety and a quenching moiety, wherein cleavage of the cleavage structure separates the fluorescing moiety from the quenching moiety, resulting in a detectable signal (e.g., FRET detection). In some embodiments, a change in quenching of signal from a donor fluorophore is detected, while in other embodiments, a change in emission from an acceptor fluorophore is detected. In

15 still other embodiments, the effect of FRET on both donor and acceptor emissions are detected.

In some embodiments of FRET detection, the fluorescence lifetime of the fluorescence emitter is measured (e.g., as in time-resolved fluorescence). While not limiting time-resolved fluorescence detection embodiments to any particular labeling systems,

20 examples of tags that are useful in time-resolved FRET measurements include europium chelate (Eu^{3+} ; Biosclair, *et al.*, J. Biomolecular Screening 5(5):319 [2000]), europium trisbipyridine cryptate (TBPEu^{3+} ; Alpha-Bazin, *et al.*, Anal. Biochem. 286(1):17 [2000]), and ruthenium ligand complex $\{[\text{Ru}(\text{bpy})_2(\text{phen-ITC})]^{2+}$; Youn, *et al.*, Anal. Biochem. 232(1):24 [1995]; Lakowicz, *et al.*, Anal. Biochem. 288:62 [2001]).

25 **c) Optimization Of Reaction Conditions**

The INVADER oligonucleotide-directed cleavage reaction is useful to detect the presence of specific nucleic acids. In addition to the considerations listed above for the selection and design of the INVADER and probe oligonucleotides, the conditions under which the reaction is to be performed may be optimized for detection of a specific target

30 sequence.

One objective in optimizing the INVADER oligonucleotide-directed cleavage assay is to allow specific detection of the fewest copies of a target nucleic acid. To achieve this end, it is desirable that the combined elements of the reaction interact with the maximum efficiency, so that the rate of the reaction (e.g., the number of cleavage events per minute) is maximized. Elements contributing to the overall efficiency of the reaction include the rate of hybridization, the rate of cleavage, and the efficiency of the release of the cleaved probe.

The rate of cleavage will be a function of the cleavage means chosen, and may be made optimal according to the manufacturer's instructions when using commercial preparations of enzymes or as described in the examples herein. The other elements (rate of hybridization, efficiency of release) depend upon the execution of the reaction, and optimization of these elements is discussed below.

Three elements of the cleavage reaction that significantly affect the rate of nucleic acid hybridization are the concentration of the nucleic acids, the temperature at which the cleavage reaction is performed and the concentration of salts and/or other charge-shielding ions in the reaction solution.

The concentrations at which oligonucleotide probes are used in assays of this type are well known in the art, and are discussed above. One example of a common approach to optimizing an oligonucleotide concentration is to choose a starting amount of oligonucleotide for pilot tests; 0.01 to 2 μ M is a concentration range used in many oligonucleotide-based assays. When initial cleavage reactions are performed, the following questions may be asked of the data: Is the reaction performed in the absence of the target nucleic acid substantially free of the cleavage product?; Is the site of cleavage specifically positioned in accordance with the design of the INVADER oligonucleotide?; Is the specific cleavage product easily detected in the presence of the uncleaved probe (or is the amount of uncut material overwhelming the chosen visualization method)?

A negative answer to any of these questions would suggest that the probe concentration is too high, and that a set of reactions using serial dilutions of the probe should be performed until the appropriate amount is identified. Once identified for a given target nucleic acid in a given sample type (e.g., purified genomic DNA, body fluid extract, lysed bacterial extract), it should not need to be re-optimized. The sample type is

important because the complexity of the material present may influence the probe concentration optimum.

Conversely, if the chosen initial probe concentration is too low, the reaction may be slow, due to inefficient hybridization. Tests with increasing quantities of the probe will identify the point at which the concentration exceeds the optimum (*e.g.*, at which it produces an undesirable effect, such as background cleavage not dependent on the target sequence, or interference with detection of the cleaved products). Since the hybridization will be facilitated by excess of probe, it is desirable, but not required, that the reaction be performed using probe concentrations just below this point.

The concentration of INVADER oligonucleotide can be chosen based on the design considerations discussed above. In some embodiments, the INVADER oligonucleotide is in excess of the probe oligonucleotide. In a preferred embodiment, the probe oligonucleotide is in excess of the INVADER oligonucleotide.

Temperature is also an important factor in the hybridization of oligonucleotides.

The range of temperature tested will depend in large part on the design of the oligonucleotides, as discussed above. Where it is desired to have a reaction be run at a particular temperature (*e.g.*, because of an enzyme requirement, for convenience, for compatibility with assay or detection apparatuses, etc.), the oligonucleotides that function in the reaction can be designed to optimally perform at the desired reaction temperature.

Each INVADER reaction includes at least two target sequence-specific oligonucleotides for the primary reaction: an upstream INVADER oligonucleotide and a downstream probe oligonucleotide. In some preferred embodiments, the INVADER oligonucleotide is designed to bind stably at the reaction temperature, while the probe is designed to freely associate and disassociate with the target strand, with cleavage occurring only

when an uncut probe hybridizes adjacent to an overlapping INVADER oligonucleotide. In preferred embodiments, the probe includes a 5' flap that is not complementary to the target, and this flap is released from the probe when cleavage occurs. The released flap can be detected directly or indirectly. In some preferred embodiments, as discussed in detail below, the released flap participate as in INVADER oligonucleotide in a secondary reaction.

Optimum conditions for the INVADER assay are generally those that allow specific detection of the smallest amount of a target nucleic acid. Such conditions may be characterized as those that yield the highest target-dependent signal in a given timeframe, or for a given amount of target nucleic acid, or that allow the highest rate of probe cleavage (*i.e.*, probes cleaved per minute).

As noted above, the concentration of the cleavage agent can affect the actual optimum temperature for a cleavage reaction. Additionally, different cleavage agents, even if used at identical concentrations, can affect reaction temperature optima differently (*e.g.*, the difference between the calculated probe T_m and the observed optimal reaction temperature may be greater for one enzyme than for another). Determination of appropriate salt corrections for reactions using different enzymes or concentrations of enzymes, or for any other variation made in reaction conditions, involves a two step process of a) measuring reaction temperature optima under the new reaction conditions, and varying the salt concentration within the T_m algorithm to produce a calculated temperature matching or closely approximating the observed optima. Measurement of an optimum reaction temperature generally involves performing reactions at a range of temperatures selected such that the range allows observation of an increase in performance as an optimal temperature is approached (either by increasing or decreasing temperatures), and a decrease in performance when an optimal temperature has been passed, thereby allowing identification of the optimal temperature or temperature range (See *e.g.*, Lyamichev, et al., *Biochemistry* 39: 9523 [2000]).

In some embodiments, a secondary reaction is used where the released cleavage fragment from a primary reaction hybridizes to a synthetic cassette to form a secondary cleavage reaction. In some preferred embodiments, the cassette comprises a fluorescing moiety and a quenching moiety, wherein cleavage of the secondary cleavage structure separates the fluorescing moiety from the quenching moiety, resulting in a detectable signal (*e.g.*, FRET detection). The secondary reaction can be configured a number of different ways. For example, in some embodiments, the synthetic cassette comprises two oligonucleotides: an oligonucleotide that contains the FRET moieties and a FRET/INVADER oligonucleotide bridging oligonucleotide that allows the INVADER oligonucleotide (*i.e.*, the released flap from the primary reaction) and the FRET

oligonucleotide to hybridize thereto, such that a cleavage structure is formed. In some embodiments, the synthetic cassette is provided as a single oligonucleotide, comprising a hairpin structure (i.e., the FRET oligonucleotide is connected at its 3' end to the bridging oligonucleotide by a loop). The loop may be nucleic acid, or a non-nucleic acid spacer or linker. The linked molecules may together be described as a FRET cassette. In the secondary reaction using a FRET cassette the released flap from the primary reaction, which acts as an INVADER oligonucleotide, should be able to associate and disassociate with the FRET cassette freely, so that one released flap can direct the cleavage of multiple FRET cassettes. It is one aspect of the assay design that all of the probe sequences may be selected to allow the primary and secondary reactions to occur at the same optimal temperature, so that the reaction steps can run simultaneously. In an alternative embodiment, the probes may be designed to operate at different optimal temperatures, so that the reaction steps are not simultaneously at their temperature optima. As noted above, the same iterative process used to select the ASR of the probe can be used in the design of the portion of the primary probe that participates in a secondary reaction.

Another determinant of hybridization efficiency is the salt concentration of the reaction. In large part, the choice of solution conditions will depend on the requirements of the cleavage agent, and for reagents obtained commercially, the manufacturer's instructions are a resource for this information. When developing an assay utilizing any particular cleavage agent, the oligonucleotide and temperature optimizations described above should be performed in the buffer conditions best suited to that cleavage agent.

In some embodiments, additional agents may be included in reaction mixtures to enhance assay performance. For example, charged compounds such as aminoglycosides and other polyamines have been used to modulate DNA and RNA conformation and function (see, *e.g.*, Earnshaw and Gait, *Nucl. Acids Res.* 26:5551 [1998]; Robinson and Wang, *Nucl. Acids Res.* 24:676 [1996]; Jerinie, *J. Mol. Biol.* 304(5):707 [2000]; Schroeder *et al.*, *EMBO* 19(1):1 [2000]). Inclusion of the aminoglycoside antibiotic neomycin sulfate (*e.g.*, at 1 μ M in a primary reaction) can enhance assay performance by, *e.g.*, reducing background signal, and therefore reducing the limit of detection of a particular INVADER assay probe set. Compounds of this type that may find use in

INVADER assay reactions include, but are not limited to, aminoglycosides, oligomerized aminoglycosides, and aminoglycoside bioconjugates, and other polyanions including, but not limited to, spermine and hexamine cobalt.

5 A "no enzyme" control allows the assessment of the stability of the labeled oligonucleotides under particular reaction conditions, or in the presence of the sample to be tested *e.g.*, in assessing the sample for contaminating nucleases). In this manner, the substrate and oligonucleotides are placed in a tube containing all reaction components, except the enzyme and treated the same as the enzyme-containing reactions. Other controls may also be included. For example, a reaction with all of the components except
10 the target nucleic acid will serve to confirm the dependence of the cleavage on the presence of the target sequence.

d) Selection of a Cleavage Agent

As demonstrated in a number of the Examples, some 5' nucleases do not require an upstream oligonucleotide to be active in a cleavage reaction. Although cleavage may
15 be slower without the upstream oligonucleotide, it may still occur (Lyamichev et al., Science 260:778 [1993], Kaiser et al., J. Biol. Chem., 274:21387 [1999]). When a DNA strand is the template or target strand to which probe oligonucleotides are hybridized, the 5' nucleases derived from DNA polymerases and some flap endonucleases (FENs), such as that from *Methanococcus jannaschii*, can cleave quite well without an upstream
20 oligonucleotide providing an overlap (Lyamichev et al., Science 260:778 [1993], Kaiser et al., J. Biol. Chem., 274:21387 [1999], and US Patent No. 5,843,669, herein incorporated by reference in its entirety). These nucleases may be selected for use in some embodiments of the INVADER assay, *e.g.*, in embodiments wherein cleavage of the probe in the absence of an INVADER oligonucleotide gives a different cleavage
25 product, which does not interfere with the intended analysis, or wherein both types of cleavage, INVADER oligonucleotide-directed and INVADER oligonucleotide-independent, are intended to occur.

In other embodiments it is preferred that cleavage of the probe be dependent on the presence of an upstream INVADER oligonucleotide, and enzyme having this
30 requirement would be used. Other FENs, such as those from *Archeaoglobus fulgidus*

(Afu) and *Pyrococcus furiosus* (Pfu), cleave an overlapped structure on a DNA target at so much greater a rate than they do a non-overlapping structure (*i.e.*, either missing the upstream oligonucleotide or having a non-overlapping upstream oligonucleotide) that they can be viewed as having an essentially absolute requirement for the overlap

(Lyamichev et al., Nat. Biotechnol., 17:292 [1999], Kaiser et al., J. Biol. Chem., 274:21387 [1999]). When an RNA target is hybridized to DNA oligonucleotide probes to form a cleavage structure, many FENs cleave the downstream DNA probe poorly, regardless of the presence of an overlap. On such an RNA-containing structure, the 5' nucleases derived from DNA polymerases have a strong requirement for the overlap, and are essentially inactive in its absence. The selection of enzymes for use in the detection of RNA targets is discussed in more detail below, in Section IV: Improved Enzymes For Use In INVADER Oligonucleotide-Directed Cleavage Reactions Comprising RNA Targets.

e) Probing For Multiple Alleles

The INVADER oligonucleotide-directed cleavage reaction is also useful in the detection and quantification of individual variants or alleles in a mixed sample population. By way of example, such a need exists in the analysis of tumor material for mutations in genes associated with cancers. Biopsy material from a tumor can have a significant complement of normal cells, so it is desirable to detect mutations even when present in fewer than 5% of the copies of the target nucleic acid in a sample. In this case, it is also desirable to measure what fraction of the population carries the mutation. Similar analyses may also be done to examine allelic variation in other gene systems, and it is not intended that the method of the present invention be limited to the analysis of tumors.

As demonstrated below, in one embodiment, reactions can be performed under conditions that prevent the cleavage of probes bearing even a single-nucleotide difference mismatch, but that permit cleavage of a similar probe that is completely complementary to the target in this region. In a preferred embodiment, a mismatch is positioned at the

nucleotide in the probe that is 5' of the site where cleavage occurs in the absence of the mismatch.

In other embodiments, the INVADER assay may be performed under conditions that have a tight requirement for an overlap (*e.g.*, using the Afu FEN for DNA target detection or the 5' nuclease of DNA polymerase for RNA target detection, as described above), providing an alternative means of detecting single nucleotide or other sequence variations. In one embodiment, the probe is selected such that the target base suspected of varying is positioned at the 5' end of the target-complementary region of this probe. The upstream INVADER oligonucleotide is positioned to provide a single base of overlap. If the target and the probe oligonucleotide are complementary at the base in question, the overlap forms and cleavage can occur. However, if the target does not complement the probe at this position, that base in the probe becomes part of a non-complementary 5' arm, no overlap between the INVADER oligonucleotide and probe oligonucleotide exists, and cleavage is suppressed.

It is also contemplated that different sequences may be detected in a single reaction. Probes specific for the different sequences may be differently labeled. For example, the probes may have different dyes or other detectable moieties, different lengths, or they may have differences in net charges of the products after cleavage. When differently labeled in one of these ways, the contribution of each specific target sequence to final product can be tallied. This has application in detecting the quantities of different versions of a gene within a mixture. Different genes in a mixture to be detected and quantified may be wild type and mutant genes (*e.g.*, as may be found in a tumor sample, such as a biopsy). In this embodiment, one might design the probes to precisely the same site, but one to match the wild-type sequence and one to match the mutant. Quantitative detection of the products of cleavage from a reaction performed for a set amount of time will reveal the ratio of the two genes in the mixture. Such analysis may also be performed on unrelated genes in a mixture. This type of analysis is not intended to be limited to two genes. Many variants within a mixture may be similarly measured.

Alternatively, different sites on a single gene may be monitored and quantified to verify the measurement of that gene. In this embodiment, the signal from each probe would be expected to be the same.

It is also contemplated that multiple probes may be used that are not differently labeled, such that the aggregate signal is measured. This may be desirable when using many probes designed to detect a single gene to boost the signal from that gene. This configuration may also be used for detecting unrelated sequences within a mix. For
5 example, in blood banking it is desirable to know if any one of a host of infectious agents is present in a sample of blood. Because the blood is discarded regardless of which agent is present, different signals on the probes would not be required in such an application of the present invention, and may actually be undesirable for reasons of confidentiality.

Just as described for the two-oligonucleotide system, above, the specificity of the
10 detection reaction will be influenced by the aggregate length of the target nucleic acid sequences involved in the hybridization of the complete set of the detection oligonucleotides. For example, there may be applications in which it is desirable to detect a single region within a complex genome. In such a case the set of oligonucleotides may be chosen to require accurate recognition by hybridization of a
15 longer segment of a target nucleic acid, often in the range of 20 to 40 nucleotides. In other instances it may be desirable to have the set of oligonucleotides interact with multiple sites within a target sample. In these cases one approach would be to use a set of oligonucleotides that recognize a smaller, and thus statistically more common, segment of target nucleic acid sequence.

20 In one preferred embodiment, the INVADER and stacker oligonucleotides may be designed to be maximally stable, so that they will remain bound to the target sequence for extended periods during the reaction. This may be accomplished through any one of a number of measures well known to those skilled in the art, such as adding extra hybridizing sequences to the length of the oligonucleotide (up to about 50 nts in total
25 length), or by using residues with reduced negative charge, such as phosphorothioates or peptide-nucleic acid residues, so that the complementary strands do not repel each other to degree that natural strands do. Such modifications may also serve to make these flanking oligonucleotides resistant to contaminating nucleases, thus further ensuring their continued presence on the target strand during the course of the reaction. In addition, the
30 INVADER and stacker oligonucleotides may be covalently attached to the target (*e.g.*, through the use of psoralen cross-linking).

f) Applications for pooled DNA and RNA samples

In some embodiments, the present invention provides methods and kits for assaying a pooled sample using INVADER detection reagents (e.g. primary probe, INVADER probe, and FRET cassette). In some preferred embodiments, the kit
5 comprises instructions on how to perform the INVADER assay and specifically how to apply the INVADER detection assay to pooled samples from many individuals, or to "pooled" samples from many cells (e.g. from a biopsy sample) from a single subject.

In particular embodiments, the present invention allows detection of polymorphisms in pooled samples combined from many individuals in a population (e.g.
10 10, 50, 100, or 500 individuals), or from a single subject where the nucleic acid sequences are from a large number of cells that are assayed at once. In this regard, the present invention allows the frequency of rare mutations in pooled samples to be detected and an allele frequency for the population established. In some embodiments, this allele frequency may then be used to statistically analyze the results of applying the INVADER
15 detection assay to an individual's frequency for the polymorphism (e.g. determined using the INVADER assay). In this regard, mutations that rely on a percent of mutants found (e.g. loss of heterozygosity mutations) may be analyzed, and the severity of disease or progression of a disease determined (See, e.g. US Patent 6,146,828 and 6,203,993 to Lapidus, hereby incorporated by reference for all purposes, where genetic testing and
20 statistical analysis are employed to find disease causing mutations or identify a patient sample as containing a disease causing mutations).

In some embodiments of the present invention, broad population screens are performed. In some preferred embodiments, pooling DNA from several hundred or a thousand individuals is optimal. In such a pool, for example, DNA from any one
25 individual would not be detectable, and any detectable signal would provide a measure of frequency of the detected allele in a broader population. The amount of DNA to be used, for example, would be set not by the number of individuals in a pool, but rather by the allele frequency to be detected. For example, in some embodiments, an assay gives ample signal from 20 to 40 ng of DNA in a 90 minute reaction. At this level of
30 sensitivity, analysis of 1 µg of DNA from a high-complexity pool would produce comparable signal from alleles present in only about 3-5% of the population.

g) Applications of RNA detection.

RNA quantitation is becoming increasingly important in basic, pharmaceutical, and clinical research. For example, quantitation of viral RNAs can predict disease progression and therapeutic efficacy. Likewise, gene expression analysis of diseased vs. normal, or untreated vs. treated, tissue can identify relevant biological responses or assess the effects of pharmacological agents. As the focus of the Human Genome Project moves toward gene expression analysis, the field will require a flexible RNA analysis technology that can quantitatively monitor multiple forms of alternatively transcribed and/or processed RNAs.

As described above for the detection of multiple alleles, multiplex formats of the RNA INVADER assay enable simultaneous expression analysis of two or more genes within the same sample. In a primary reaction, one-nucleotide overlap-substrates are generated by the hybridization of INVADER oligonucleotides and probe oligonucleotides to their respective RNA targets. Each probe contains a specific, target-complementary region and a distinctive non-complementary 5' flap that is associated only with that specific mRNA in that assay. The distinctive flaps may be distinguished in any of the myriad ways disclosed herein (*e.g.*, with different labels, different secondary cleavage systems having different labels, specific antibodies, different sizes when resolved, different sequences detected by hybridization in solution or on surfaces, etc.)

While the RNA invasive cleavage assay, like the method used for DNA detection described above, can use two invasive cleavage reactions in sequence (described below, in Section II of the Detailed Description of the Invention), its preference for the 5' nucleases derived from DNA polymerases (described in detail in Section IV of the Detailed Description of the Invention) indicates that additional format changes are preferred. Unlike the FEN 5' nucleases generally used for detection of DNA targets, optimal signal amplification with the DNA Pol-related 5' nucleases occurs only when a probe turnover mechanism is employed in both the primary and secondary reactions (in contrast to an INVADER oligonucleotide turnover mechanism, wherein an INVADER oligonucleotide cycles, *e.g.*, to direct the cleavage of multiple FRET cassettes, as

described below, in Section II of this Detailed Description). Consequently, in preferred embodiments, RNA detection uses sequential operation of the two reactions, rather than simultaneous reaction performance. Because the reactions are performed truly sequentially, in these embodiments, the RNA INVADER assay signal accumulates linearly in both a target- and time-dependent manner. In contrast, the primary and secondary reactions of the DNA INVADER assay, when run concurrently, amplify signal as a linear function of target level, but as a quadratic function of time. In the sequential embodiments, the RNA INVADER assay uses two separate oligonucleotides, a secondary probe (*e.g.*, a FRET probe) and secondary target, for signal generation.

A key feature of the RNA invasive cleavage assay is its ability to discriminate highly homologous RNA sequences, such as those found in cytochrome P450 gene families. Like the DNA INVADER assay, the RNA INVADER assay can discriminate single-base changes. In some embodiments, the first 5' complementary base of each probe is positioned at a non-conserved site in its mRNA target, so that a mismatch prevents formation of the overlap-structure, and thus prevents cleavage of the probe. Alternatively spliced mRNA variants can be specifically detected by positioning the cleavage site at a splice junction.

To monitor large changes in mRNA levels, the dynamic range of the assay can be extended using real-time analysis. However, since the assay generates signal linearly with time or target level, simply varying the amount of sample added per reaction and calculating the copies of mRNA per ng total RNA enables accurate quantitation with a single endpoint measurement on low-cost instrumentation. Further, in cases where absolute quantitation is not necessary, the assay's linear signal amplification mechanism and reproducibility also eliminate the need for a standard curve and enable simple and precise relative quantitation of any one gene.

The RNA INVADER assay is particularly suited for detecting alternatively spliced or edited RNA variants because even a single base change at the overlap site affects 5' nuclease cleavage. All areas requiring RNA quantitation, such as high-throughput screening in drug discovery research, monitoring drug metabolism and safety in clinical trials, and clinical load monitoring of viral RNA can use this technology.

Splice variants can be monitored in at least two ways with the assay: 1) detection of an individual exon or 2) detection of a specific splice junction.

To examine an RNA population for variants having more or fewer exons after splicing, INVADER assay probe sets are designed for each of the exons of interest (or for all exons in the mature RNA). Quantitation of exons, independent of how many mRNAs they reside in, may provide information about the number of splice variants for a given gene, as well as indicate the levels of expression for each exon. Mini *in vitro* transcripts containing only one or a few exons can be generated for each probe set so that absolute quantitation can be performed for each exon, thus enabling accurate comparisons of exon levels. If it is known that a particular exon is present in all known variants, in some embodiments, a probe set is designed for that exon for use as an internal control to normalize across different samples. RNAs having a one copy of each exon (*e.g.*, "normally spliced" RNA) should produce signal from the collection of probe sets in certain relative amounts (which should be essentially equal for all exons, corrected for variations in the sensitivity of individual probe sets; see Section V). Alterations in splicing alter the relative amounts of the exons. For example, if all of the produced RNAs are missing one of the normal exons, the signal for that exon drops toward zero, while if half of the RNAs are missing that exon, the signal for that exon drops toward 50%. More complex combinations of splice variations and mixtures of differently spliced mRNAs yield more complex and more informative profiles. Detection is not limited to exons. RNA populations may also be monitored for the presence of intron sequences that are usually removed by splicing. Such global exon screening provides biologically relevant or diagnostic information when comparing normal vs. diseased tissue or untreated vs. treated cells (*e.g.*, in pharmacogenomic screening assays). An array-based description of this type of measurement is referred to as alternative splicing detector arrays (D. Black, Cell 103:367 [2000]). It is contemplated that the mRNA INVADER assay gives similar results but with greater specificity and more accurate quantitation than the oligonucleotide array formats.

In an alternative embodiment, alternatively spliced mRNAs is detected by examination of specific splice junctions. The advantage of monitoring the splice sites, as

opposed to the exons themselves, is that even splice variants involving very small exons (e.g. < 10 nts) are accurately detected with the assay.

Additionally, in some embodiments, the mRNA INVADER assay is also used to monitor alternative start and stop sites in the mRNA, and is used to monitor lifetimes of processed and unprocessed RNAs and RNA fragments (e.g., as used in timecourse studies following induction).

II. Signal Enhancement By Incorporating The Products Of An Invasive Cleavage Reaction Into A Subsequent Invasive Cleavage Reaction

As noted above, the oligonucleotide product released by the invasive cleavage can be used subsequently in any reaction or read-out method that uses oligonucleotides in the size range of a cleavage product. In addition to the reactions involving primer extension and transcription, described herein, another enzymatic reaction that makes use of oligonucleotides is the invasive cleavage reaction. The present invention provides means of using the oligonucleotide released in a primary invasive cleavage reaction as a component to complete a cleavage structure to enable a secondary invasive cleavage reaction. It is not intended that the sequential use of the invasive cleavage product be limited to a single additional step. It is contemplated that many distinct invasive cleavage reactions may be performed in sequence.

The polymerase chain reaction uses a DNA replication method to create copies of a targeted segment of nucleic acid at a logarithmic rate of accumulation. This is made possible by the fact that when the strands of DNA are separated, each individual strand contains sufficient information to allow assembly of a new complementary strand. When the new strands are synthesized the number of identical molecules has doubled. Within 20 iterations of this process, the original may be copied 1 million-fold, making very rare sequences easily detectable. The mathematical power of a doubling reaction has been incorporated into a number of amplification assays.

By performing multiple, sequential invasive cleavage reactions the method of the present invention captures an exponential mathematical advantage without producing additional copies of the target analyte. In a simple invasive cleavage reaction the yield,

Y, is simply the turnover rate, K, multiplied by the time of the reaction, t (*i.e.*, $Y = (K)(t)$). If Y is used to represent the yield of a simple reaction, then the yield of a compound (*i.e.*, a multiple, sequential reaction), assuming that each of the individual invasive cleavage steps has the same turnover rate, can be simply represented as Y^n , where n is the number of invasive cleavage reactions that have been performed in the series. If the yields of each step differ the ultimate yield can be represented as the product of the multiplication of the yields of each individual reaction in the series. For example, if a primary invasive cleavage reaction can produce one thousand products in 30 minutes, and each of those products can in turn participate in 1000 additional reactions, there will be 1000^2 copies (1000×1000) of the ultimate product in a second reaction. If a third reaction is added to the series, then the theoretical yield will be 1000^3 ($1000 \times 1000 \times 1000$). In the methods of the present invention the exponent comes from the number of invasive cleavage reactions in the cascade. This can be contrasted to the amplification methods described above (*e.g.*, PCR) in which Y is limited to 2 by the number of strands in duplex DNA, and the exponent n is the number of cycles performed, so that many iterations are necessary to accumulate large amounts of product.

To distinguish the exponential amplifications described above from those of the present invention, the former can be considered reciprocating reactions because the products the reaction feed back into the same reaction (*e.g.*, event one leads to some number of events 2, and each event 2 leads back to some number of events 1). In contrast, the events in some embodiments of the present invention are sequential (*e.g.*, event 1 leads to some number of events 2; each event 2 leads to some number of events 3, etc., and no event can contribute to an event earlier in the chain).

The sensitivity of the reciprocating methods is also one of the greatest weaknesses when these assays are used to determine if a target nucleic acid sequence is present or absent in a sample. Because the product of these reactions is detectable copies of the starting material, contamination of a new reaction with the products of an earlier reaction can lead to false positive results, (*i.e.*, the apparent detection of the target nucleic acid in samples that do not actually contain any of that target analyte). Furthermore, because the concentration of the product in each positive reaction is so high, amounts of DNA sufficient to create a strong false positive signal can be communicated to new reactions

very easily either by contact with contaminated instruments or by aerosol. In contrast to the reciprocating methods, the most concentrated product of the sequential reaction (*i.e.*, the product released in the ultimate invasive cleavage event) is not capable of initiating a like reaction or cascade if carried over to a fresh test sample. This is a marked advantage over the exponential amplification methods described above because the reactions of the present invention may be performed without the costly containment arrangements (*e.g.*, either by specialized instruments or by separate laboratory space) required by any reciprocating reaction. While the products of a penultimate event may be inadvertently transferred to produce a background of the ultimate product in the absence of the a target analyte, the contamination would need to be of much greater volume to give an equivalent risk of a false positive result.

When the term sequential is used it is not intended to limit the invention to configurations in which that one invasive cleavage reaction or assay must be completed before the initiation of a subsequent reaction for invasive cleavage of a different probe. Rather, the term refers to the order of events as would occur if only single copies of each of the oligonucleotide species were used in an assay. The primary invasive cleavage reaction refers to that which occurs first, in response to the formation of the cleavage structure on the target nucleic acid. Subsequent reactions may be referred to as secondary, tertiary and so forth, and may involve artificial "target" strands that serve only to support assembly of a cleavage structure, and which are unrelated to the nucleic acid analyte of interest. While the complete assay may, if desired, be configured with each step of invasive cleavage separated either in space (*e.g.*, in different reaction vessels) or in time (*e.g.*, using a shift in reaction conditions, such as temperature, enzyme identity or solution condition, to enable the later cleavage events), it is also contemplated that all of the reaction components may be mixed so that secondary reactions may be initiated as soon as product from a primary cleavage becomes available. In such a format, primary, secondary and subsequent cleavage events involving different copies of the cleavage structures may take place simultaneously.

Several levels of this sort of linear amplification can be envisioned, in which each successive round of cleavage produces an oligonucleotide that can participate in the cleavage of a different probe in subsequent rounds. The primary reaction would be

specific for the analyte of interest with secondary (and tertiary, etc.) reactions being used to generate signal while still being dependent on the primary reaction for initiation.

The released product may perform in several capacities in the subsequent reactions. For example, the product of one invasive cleavage reaction becomes the
5 INVADER oligonucleotide to direct the specific cleavage of another probe in a second reaction. In such an example, the first invasive cleavage structure is formed by the annealing of the INVADER oligonucleotide and the probe oligonucleotide (Probe 1) to the first target nucleic acid (Target 1). The target nucleic acid is divided into three regions based upon which portions of the INVADER and probe oligonucleotides are
10 capable of hybridizing to the target. Region 1 of the target has complementarity to only the INVADER oligonucleotide; region 3 of the target has complementarity to only the probe; and region 2 of the target has an overlap between the INVADER and probe oligonucleotides.

Cleavage of Probe 1 releases the "Cut Probe 1". The released Probe 1 is then
15 used as the INVADER oligonucleotide in second cleavage. The second cleavage structure is formed by the annealing of the Cut Probe 1, a second probe oligonucleotide ("Probe 2") and a second target nucleic acid ("Target 2"). In some embodiments, Probe 2 and the second target nucleic acid are covalently connected, preferably at their 3' and 5' ends, respectively, thus forming a hairpin stem and loop, termed herein a "cassette". The
20 loop may be nucleic acid or a non-nucleic acid spacer or linker. Inclusion of an excess of the cassette molecule allows each Cut Probe 1 to serve as an INVADER to direct the cleavage of multiple copies of the cassette.

Probe 2 may be labeled and detection of cleavage of the second cleavage structure may be accomplished by detecting the labeled cut Probe 2; the label may a radioisotope
25 (*e.g.*, ^{32}P , ^{35}S), a fluorophore (*e.g.*, fluorescein), a reactive group capable of detection by a secondary agent (*e.g.*, biotin/streptavidin), a positively charged adduct which permits detection by selective charge reversal (as discussed in Section IV above), etc.

Alternatively, the cut Probe 2 may be used in a tailing reaction, or to complete or activate a protein-binding site, or may be detected or used by any of the means for detecting or
30 using an oligonucleotide described herein.

In other embodiments, probe oligonucleotides that are cleaved in the primary reaction can be designed to fold back on themselves (*i.e.*, they contain a region of self-complementarity) to create a molecule that can serve as both the INVADER and target oligonucleotide (termed here an "IT" complex). The IT complex then enables cleavage of a different probe present in the secondary reaction. Inclusion of an excess of the secondary probe molecule ("Probe 2"), allows each IT molecule to serve as the platform for the generation of multiple copies of cleaved secondary probe. The target nucleic acid is divided into three regions based upon which portions of the INVADER and probe oligonucleotides are capable of hybridizing to the target (as discussed above and it is noted that the target may be divided into four regions if a stacker oligonucleotide is employed). The second cleavage structure is formed by the annealing of the second probe ("Probe 2") to the fragment of Probe 1 ("Cut Probe 1") that was released by cleavage of the first cleavage structure. The Cut Probe 1 forms a hairpin or stem/loop structure near its 3' terminus by virtue of the annealing of the regions of self-complementarity contained within Cut Probe 1 (this self-annealed Cut Probe 1 forms the IT complex). The IT complex (Cut Probe 1) is divided into three regions. Region 1 of the IT complex has complementarity to the 3' portion of Probe 2; region 2 has complementarity to both the 3' end of Cut Probe 1 and to the 5' portion of Probe 2; and region 3 contains the region of self-complementarity (*i.e.*, region 3 is complementary to the 3' portion of the Cut Probe 1). Note that with regard to the IT complex (*i.e.*, Cut Probe 1), region 1 is located upstream of region 2 and region 2 is located upstream of region 3. As for other embodiments of invasive cleavage, region "2" can represent a region where there is a physical, but not sequence, overlap between the INVADER oligonucleotide portion of the Cut Probe 1 and the Probe 2 oligonucleotide.

The cleavage products of the secondary invasive cleavage reaction (*i.e.*, Cut Probe 2) can either be detected, or can in turn be designed to constitute yet another integrated INVADER-target complex to be used with a third probe molecule, again unrelated to the preceding targets.

It is envisioned that the oligonucleotide product of a primary cleavage reaction may fill the role of any of the oligonucleotides described herein (*e.g.*, it may serve as a target strand without an attached INVADER oligonucleotide-like sequence, or it may

serve as a stacker oligonucleotide, as described above), to enhance the turnover rate seen in the secondary reaction by stabilizing the probe hybridization through coaxial stacking.

Secondary cleavage reactions in some preferred embodiments of the present invention include the use of FRET cassettes. Such molecules provide both a secondary target and a FRET labeled cleavable sequence, allowing homogeneous detection (*i.e.*, without product separation or other manipulation after the reaction) of the sequential invasive cleavage reaction. Other preferred embodiments use a secondary reaction system in which the FRET probe and synthetic target are provided as separate oligonucleotides.

In a preferred embodiment, each subsequent reaction is facilitated by (*i.e.*, is dependent upon) the product of the previous cleavage, so that the presence of the ultimate product may serve as an indicator of the presence of the target analyte. However, cleavage in the second reaction need not be dependent upon the presence of the product of the primary cleavage reaction; the product of the primary cleavage reaction may merely measurably enhance the rate of the second cleavage reaction.

In summary, the INVADER assay cascade (*i.e.*, sequential invasive cleavage reactions) of the present invention is a combination of two or more linear assays that allows the accumulation of the ultimate product at an exponential rate, but without significant risk of carryover contamination. It is important to note that background that does not arise from sequential cleavage, such as thermal breakage of the secondary probe, generally increases linearly with time. In contrast, signal generation from a 2-step sequential reaction follows quadratic kinetics. Thus, collection of data as a time course, either by taking time points or through the use of an instrument that allows real-time detection during the INVADER assay reaction incubations, provides the attractive capability of discriminating between the true signal and any background solely on the basis of quadratic versus linear increases in signal over time. For example, when viewed graphically, the real signal will appear as a quadratic curve, while any accumulating background will be linear, and thus easy to distinguish, even if the absolute level of the background signal (*e.g.*, fluorescence in a FRET detection format) is substantial.

The sequential invasive cleavage amplification of the present invention can be used as an intermediate boost to any of the detection methods (*e.g.*, gel based analysis by

either standard or by charge reversal), polymerase tailing, and incorporation into a protein binding region, described herein. When used in such combinations the increased production of a specific cleavage product in the invasive cleavage assay reduces the burdens of sensitivity and specificity on the read-out systems, thus facilitating their use.

5 In addition to enabling a variety of detection platforms, the cascade strategy is suitable for multiplex analysis of individual analytes (*i.e.*, individual target nucleic acids) in a single reaction. The multiplex format can be categorized into two types. In one case, it is desirable to know the identity (and amount) of each of the analytes that can be present in a clinical sample, or the identity of each of the analytes as well as an internal
10 control. To identify the presence of multiple individual analytes in a single sample, several distinct secondary amplification systems may be included. Each probe cleaved in response to the presence of a particular target sequence (or internal control) can be designed to trigger a different cascade coupled to different detectable moieties, such as different sequences to be extended by DNA polymerase or different dyes in an FRET
15 format. The contribution of each specific target sequence to final product can thereby be tallied, allowing quantitative detection of different genes or alleles in a sample containing a mixture of genes or alleles.

In the second configuration, it is desirable to determine if any of several analytes are present in a sample, but the exact identity of each does not need to be known. For
20 example, in blood banking it is desirable to know if any one of a host of infectious agents is present in a sample of blood. Because the blood is discarded regardless of which agent is present, different signals on the probes would not be required in such an application of the present invention, and may actually be undesirable for reasons of confidentiality. In this case, the 5' arms (*i.e.*, the 5' portion that will be released upon cleavage) of the
25 different analyte-specific probes would be identical and would therefore trigger the same secondary signal cascade. A similar configuration would permit multiple probes complementary to a single gene to be used to boost the signal from that gene or to ensure inclusivity when there are numerous alleles of a gene to be detected.

In the primary INVADER assay reaction, there are two potential sources of
30 background. The first is from INVADER oligonucleotide-independent cleavage of probe annealed to the target, to itself, or to one of the other oligonucleotides present in the

reaction. The use of an enzyme that cannot efficiently cleave a structure that lacks a primer is preferred for this reason. The enzyme *Pfu* FEN-1 gives no detectable cleavage in the absence of the upstream oligonucleotide or even in the presence of an upstream oligonucleotide that fails to invade the probe-target complex. This indicates that the *Pfu* FEN-1 endonuclease is a suitable enzyme for use in the methods of the present invention.

Other structure-specific nucleases may be suitable as well. As discussed in the first example, some 5' nucleases can be used in conditions that significantly reduce this primer-independent cleavage. For example, it has been shown that when the 5' nuclease of DNAPtaq is used to cleave hairpins the primer-independent cleavage is markedly reduced by the inclusion of a monovalent salt in the reaction (Lyamichev, *et al.*, [1993], *supra*).

III. Effect of ARRESTOR Molecules on Signal and Background in Sequential Invasive Cleavage Reactions.

As described above, the concentration of the probe that is cleaved can be used to increase the rate of signal accumulation, with higher concentrations of probe yielding higher final signal. However, the presence of large amounts of residual uncleaved probe can present problems for subsequent use of the cleaved products for detection or for further amplification. If the subsequent step is a simple detection (*e.g.*, by gel resolution), the excess uncut material may cause background by streaking or scattering of signal, or by overwhelming a detector (*e.g.*, over-exposing a film in the case of radioactivity, or exceeding the quantitative detection limits of a fluorescence imager).

This can be overcome by partitioning the product from the uncut probe. In more complex detection methods, the cleaved product may be intended to interact with another entity to indicate cleavage. As noted above, the cleaved product can be used in any reaction that makes use of oligonucleotides, such as hybridization, primer extension, ligation, or the direction of invasive cleavage. In each of these cases, the fate of the residual uncut probe should be considered in the design of the reaction. In a primer extension reaction, the uncut probe can hybridize to a template for extension. If cleavage is required to reveal the correct 3' end for extension, the hybridized uncut probe will not be extended. It may, however, compete with the cleaved product for the template. If the template is in excess

of the combination of cleaved and uncleaved probe, then both of the latter should be able to find a copy of template for binding. If, however, the template is limiting, any competition may reduce the portion of the cleaved probe that can find successfully bind to the available template. If a vast excess of probe was used to drive the initial reaction, the remainder may also be in vast excess over the cleavage product, and thus may provide a very effective competitor, thereby reducing the amount of the final reaction (*e.g.*, extension) product for ultimate detection.

The participation of the uncut probe material in a secondary reaction can also contribute to background in these reactions. While the presentation of a cleaved probe for a subsequent reaction may represent an ideal substrate for the enzyme to be used in the next step, some enzymes may also be able to act, albeit inefficiently, on the uncut probe as well. It was shown during the development of the present invention that transcription can be promoted from a nicked promoter even when one side of the nick has additional unpaired nucleotides. Similarly, when the subsequent reaction is to be an invasive cleavage, the uncleaved probe may bind to the elements intended to form the second cleavage structure with the cleaved probe. In experiments conducted during the development of the present invention, it was found that some of the 5' nucleases described herein can catalyze some measure of cleavage of defective structures. Even at a low level, this aberrant cleavage can be misinterpreted as positive target-specific cleavage signal.

With these negative effects of the surfeit of uncut probe considered, there is clearly a need for some method of preventing these interactions. As noted above, it is possible to partition the cleaved product from the uncut probe after the primary reaction by traditional methods. However, these methods are often time consuming, may be expensive (*e.g.*, disposable columns, gels, etc.), and may increase the risk for sample mishandling or contamination. It is far preferable to configure the sequential reactions such that the original sample need not be removed to a new vessel for subsequent reaction.

The present invention provides a method for reducing interactions between the primary probe and other reactants. This method provides a means of specifically diverting the uncleaved probes from participation in the subsequent reactions. The

diversion is accomplished by the inclusion in the next reaction step an agent designed to specifically interact with the uncleaved primary probe. While the primary probe in an invasive cleavage reaction is discussed for reasons of convenience, it is contemplated that the ARRESTOR molecules may be used at any reaction step within a chain of invasive
5 cleavage steps, as needed or desired for the design of an assay. It is not intended that the ARRESTOR molecules of the present invention be limited to any particular step.

The method of diverting the residual uncut probes from a primary reaction makes use of agents that can be specifically designed or selected to bind to the uncleaved probe molecules with greater affinity than to the cleaved probes, thereby allowing the cleaved
10 probe species to effectively compete for the elements of the subsequent reaction, even when the uncut probe is present in vast excess. These agents have been termed "ARRESTOR molecules," due to their function of stopping or arresting the primary probe from participation in the later reaction. In various Examples below, an oligonucleotide is provided as an ARRESTOR molecule in an invasive cleavage assay. It
15 can be appreciated that any molecule or chemical that can discriminate between the full-length uncut probe and the cleaved probe, and that can bind or otherwise disable the uncleaved probe preferentially may be configured to act as ARRESTOR molecules within the meaning of the present invention. For example, antibodies can be derived with such specificity, as can the "aptamers" that can be selected through multiple steps of *in*
20 *vitro* amplification (*e.g.*, "SELEX," U.S. Patent Nos. 5,270,163 and 5,567,588; herein incorporated by reference) and specific rounds of capture or other selection means.

In one embodiment, the ARRESTOR molecule is an oligonucleotide. In another embodiment the ARRESTOR oligonucleotides is a composite oligonucleotide, comprising two or more short oligonucleotides that are not covalently linked, but that
25 bind cooperatively and are stabilized by co-axial stacking. In a preferred embodiment, the oligonucleotide is modified to reduce interactions with the cleavage agents of the present invention. When an oligonucleotide is used as an ARRESTOR oligonucleotide, it is intended that it not participate in the subsequent reactive step. The binding of the ARRESTOR oligonucleotide to the primary probe may, either with the participation of
30 the secondary target, or without such participation, create a bifurcated structure that is a substrate for cleavage by the 5' nucleases used in some embodiments of the methods of

some portion of these cleavage products, thereby reducing the absolute level of signal generated from a given amount of target material. If this same ARRESTOR oligonucleotide has the simultaneous effect of reducing the background of the reaction (*i.e.*, from non-target specific cleavage) by a factor that is greater than the factor of reduction in the specific signal, then the significance of the signal (*i.e.*, the ratio of signal to background), is increased, even with the lower amount of absolute signal. Any potential ARRESTOR molecule design may be tested in a simple fashion by comparing the levels of background and specific signals from reactions that lack ARRESTOR molecules to the levels of background and specific signal from similar reactions that include ARRESTOR oligonucleotides. What constitutes an acceptable level of tradeoff of absolute signal for specificity will vary for different applications (*e.g.*, target levels, read-out sensitivity, etc.), and can be determined by any individual user using the methods of the present invention.

IV. Improved Enzymes For Use In INVADER Oligonucleotide-Directed Cleavage Reactions Comprising RNA Targets;

A cleavage structure is defined herein as a structure that is formed by the interaction of a probe oligonucleotide and a target nucleic acid to form a duplex, the resulting structure being cleavable by a cleavage agent, including but not limited to an enzyme. The cleavage structure is further defined as a substrate for specific cleavage by the cleavage means in contrast to a nucleic acid molecule that is a substrate for nonspecific cleavage by agents such as phosphodiesterases. In considering improvements to enzymatic cleavage agents, one may consider the action of said enzymes on any structures that fall within the definition of a cleavage structure. Specific cleavage at any site within such a structure is contemplated.

Improvements in an enzyme may be an increased or decreased rate of cleavage of one or more types of structures. Improvements may also result in more or fewer sites of cleavage on one or more of said cleavage structures. In developing a library of new structure-specific nucleases for use in nucleic acid cleavage assays, improvements may

have many different embodiments, each related to the specific substrate structure used in a particular assay.

As an example, one embodiment of the INVADER oligonucleotide-directed cleavage assay of the present invention may be considered. In the INVADER oligonucleotide-directed cleavage assay, the accumulation of cleaved material is influenced by several features of the enzyme behavior. Not surprisingly, the turnover rate, or the number of structures that can be cleaved by a single enzyme molecule in a set amount of time, is very important in determining the amount of material processed during the course of an assay reaction. If an enzyme takes a long time to recognize a substrate (e.g., if it is presented with a less-than-optimal structure), or if it takes a long time to execute cleavage, the rate of product accumulation is lower than if these steps proceeded quickly. If these steps are quick, yet the enzyme "holds on" to the cleaved structure, and does not immediately proceed to another uncut structure, the rate will be negatively affected.

Enzyme turnover is not the only way in which enzyme behavior can negatively affect the rate of accumulation of product. When the means used to visualize or measure product is specific for a precisely defined product, products that deviate from that definition may escape detection, and thus the rate of product accumulation may appear to be lower than it is. For example, if one had a sensitive detector for trinucleotides that could not see di- or tetranucleotides, or any sized oligonucleotide other than 3 residues, in the INVADER-directed cleavage assay of the present invention any errant cleavage would reduce the detectable signal proportionally. It can be seen from the cleavage data presented here that, while there is usually one site within a probe that is favored for cleavage, there are often products that arise from cleavage one or more nucleotides away from the primary cleavage site. These are products that are target-dependent, and are thus not non-specific background. Nevertheless, if a subsequent visualization system can detect only the primary product, these represent a loss of signal. One example of such a selective visualization system is the charge reversal readout presented herein, in which the balance of positive and negative charges determines the behavior of the products. In such a system the presence of an extra nucleotide or the absence of an expected nucleotide can exclude a legitimate cleavage product from ultimate detection by leaving

that product with the wrong balance of charge. It can be easily seen that any assay that can sensitively distinguish the nucleotide content of an oligonucleotide, such as standard stringent hybridization, suffers in sensitivity when some fraction of the legitimate product is not eligible for successful detection by that assay.

5 These discussions suggest two highly desirable traits in any enzyme to be used in the method of the present invention. First, the more rapidly the enzyme executes an entire cleavage reaction, including recognition, cleavage and release, the more signal it may potentially created in the INVADER oligonucleotide-directed cleavage assay. Second, the more successful an enzyme is at focusing on a single cleavage site within a
10 structure, the more of the cleavage product can be successfully detected in a selective read-out.

 The rationale cited above for making improvements in enzymes to be used in the INVADER oligonucleotide-directed cleavage assay are meant to serve as an example of one direction in which improvements might be sought, but not as a limit on either the
15 nature or the applications of improved enzyme activities. As another direction of activity change that would be appropriately considered improvement, the DNAP-associated 5' nucleases may be used as an example. In creating some of the polymerase-deficient 5' nucleases described herein it was found that the those that were created by deletion of substantial portions of the polymerase domain, assumed activities that were weak or
20 absent in the parent proteins. These activities included the ability to cleave non-forked structures, a greatly enhanced ability to exonucleolytically remove nucleotides from the 5' ends of duplexed strands, and a nascent ability to cleave circular molecules without benefit of a free 5' end.

 In addition to the 5' nucleases derived from DNA polymerases, the present
25 invention also contemplates the use of structure-specific nucleases that are not derived from DNA polymerases. For example, a class of eukaryotic and archaebacterial endonucleases have been identified which have a similar substrate specificity to 5' nucleases of Pol I-type DNA polymerases. These are the FEN1 (Flap EndoNuclease), RAD2, and XPG (Xeroderma Pigmentosa-complementation group G) proteins. These
30 proteins are involved in DNA repair, and have been shown to favor the cleavage of structures that resemble a 5' arm that has been displaced by an extending primer during

polymerization. Similar DNA repair enzymes have been isolated from single cell and higher eukaryotes and from archaea, and there are related DNA repair proteins in eubacteria. Similar 5' nucleases have also been associated with bacteriophage such as T5 and T7.

5 Recently, the 3-dimensional structures of DNAP^{Taq} and T5 phage 5'-exonuclease were determined by X-ray diffraction (Kim *et al.*, Nature 376:612 [1995]; and Ceska *et al.*, Nature 382:90 [1995]). The two enzymes have very similar 3-dimensional structures despite limited amino acid sequence similarity. The most striking feature of the T5 5'-exonuclease structure is the existence of a triangular hole formed by the active site of
10 the protein and two alpha helices. This same region of DNAP^{Taq} is disordered in the crystal structure, indicating that this region is flexible, and thus is not shown in the published 3-dimensional structure. However, the 5' nuclease domain of DNAP^{Taq} is likely to have the same structure, based its overall 3-dimensional similarity to T5 5'-exonuclease, and that the amino acids in the disordered region of the DNAP^{Taq}
15 protein are those associated with alpha helix formation. The existence of such a hole or groove in the 5' nuclease domain of DNAP^{Taq} was predicted based on its substrate specificity (Lyamichev *et al.*, *supra*).

 It has been suggested that the 5' arm of a cleavage structure must thread through the helical arch described above to position said structure correctly for cleavage (Ceska *et al.*, *supra*).
20 One of the modifications of 5' nucleases described herein opened up the helical arch portion of the protein to allow improved cleavage of structures that cut poorly or not at all (*e.g.*, structures on circular DNA targets that would preclude such threading of a 5' arm). The gene construct that was chosen as a model to test this approach was the one called CLEAVASE BN, which was derived from DNAP^{Taq} but
25 does not contain the polymerase domain. It comprises the entire 5' nuclease domain of DNAP^{Taq}, and thus should be very close in structure to the T5 5' exonuclease. This 5' nuclease was chosen to demonstrate the principle of such a physical modification on proteins of this type. The arch-opening modification of the present invention is not intended to be limited to the 5' nuclease domains of DNA polymerases, and is
30 contemplated for use on any structure-specific nuclease that includes such an aperture as a limitation on cleavage activity. The present invention contemplates the insertion of a

thrombin cleavage site into the helical arch of DNAPs derived from the genus *Thermus* as well as 5' nucleases derived from DNAPs derived from the genus *Thermus*. The specific example shown herein using the CLEAVASE BN/thrombin nuclease merely illustrates the concept of opening the helical arch located within a nuclease domain. As the amino acid sequence of DNAPs derived from the genus *Thermus* are highly conserved, the teachings of the present invention enable the insertion of a thrombin site into the helical arch present in these DNAPs and 5' nucleases derived from these DNAPs.

The opening of the helical arch was accomplished by insertion of a protease site in the arch. This allowed post-translational digestion of the expressed protein with the appropriate protease to open the arch at its apex. Proteases of this type recognize short stretches of specific amino acid sequence. Such proteases include thrombin and factor Xa. Cleavage of a protein with such a protease depends on both the presence of that site in the amino acid sequence of the protein and the accessibility of that site on the folded intact protein. Even with a crystal structure it can be difficult to predict the susceptibility of any particular region of a protein to protease cleavage. Absent a crystal structure it must be determined empirically.

In selecting a protease for a site-specific cleavage of a protein that has been modified to contain a protease cleavage site, a first step is to test the unmodified protein for cleavage at alternative sites. For example, DNAPTaq and CLEAVASE BN nuclease were both incubated under protease cleavage conditions with factor Xa and thrombin proteases. Both nuclease proteins were cut with factor Xa within the 5' nuclease domain, but neither nuclease was digested with large amounts of thrombin. Thus, thrombin was chosen for initial tests on opening the arch of the CLEAVASE BN enzyme.

In the protease/ CLEAVASE modifications described herein the factor Xa protease cleaved strongly in an unacceptable position in the unmodified nuclease protein, in a region likely to compromise the activity of the end product. Other unmodified nucleases contemplated herein may not be sensitive to the factor Xa, but may be sensitive to thrombin or other such proteases. Alternatively, they may be sensitive to these or other such proteases at sites that are immaterial to the function of the nuclease sought to be modified. In approaching any protein for modification by addition of a protease cleavage site, the unmodified protein should be tested with the proteases under

consideration to determine which proteases give acceptable levels of cleavage in other regions.

Working with the cloned segment of DNAPTaq from which the CLEAVASE BN protein is expressed, nucleotides encoding a thrombin cleavage site were introduced in-frame near the sequence encoding amino acid 90 of the nuclease gene. This position was determined to be at or near the apex of the helical arch by reference to both the 3-dimensional structure of DNAPTaq, and the structure of T5 5' exonuclease. The encoded amino acid sequence, LVPRGS, was inserted into the apex of the helical arch by site-directed mutagenesis of the nuclease gene. The proline (P) in the thrombin cleavage site was positioned to replace a proline normally in this position in CLEAVASE BN because proline is an alpha helix-breaking amino acid, and may be important for the 3-dimensional structure of this arch. This construct was expressed, purified and then digested with thrombin. The digested enzyme was tested for its ability to cleave a target nucleic acid, bacteriophage M13 genomic DNA, that does not provide free 5' ends to facilitate cleavage by the threading model.

While the helical arch in this nuclease was opened by protease cleavage, it is contemplated that a number of other techniques could be used to achieve the same end. For example, the nucleotide sequence could be rearranged such that, upon expression, the resulting protein would be configured so that the top of the helical arch (amino acid 90) would be at the amino terminus of the protein, the natural carboxyl and amino termini of the protein sequence would be joined, and the new carboxyl terminus would lie at natural amino acid 89. This approach has the benefit that no foreign sequences are introduced and the enzyme is a single amino acid chain, and thus may be more stable than the cleaved 5' nuclease. In the crystal structure of DNAPTaq, the amino and carboxyl termini of the 5'-exonuclease domain lie in close proximity to each other, which suggests that the ends may be directly joined without the use of a flexible linker peptide sequence as is sometimes necessary. Such a rearrangement of the gene, with subsequent cloning and expression could be accomplished by standard PCR recombination and cloning techniques known to those skilled in the art.

The INVADER invasive cleavage reaction has been shown to be useful in the detection of RNA target strands (*See e.g.*, U.S. Patent 6,001,567, incorporated herein by

reference in its entirety). As with the INVADER assay for the detection of DNA (Lyamichev *et al.*, Nat. Biotechnol., 17:292 [1999]), the reactions may be run under conditions that permit the cleavage of many copies of a probe for each copy of the target RNA present in the reaction. In one embodiment, the reaction may be performed at a temperature close to the melting temperature (T_m) of the probe that is cleaved, such that the cleaved and uncleaved probes readily cycle on and off the target strand without temperature cycling. Each time a full-length probe binds to the target in the presence of the INVADER oligonucleotide, it may be cleaved by a 5' nuclease enzyme, resulting in an accumulation of cleavage product. The accumulation is highly specific for the sequence being detected, and may be configured to be proportional to both time and target concentration of the reaction. In another embodiment, the temperature of the reaction may be shifted (*i.e.*, it may be raised to a temperature that will cause the probe to dissociate) then lowered to a temperature at which a new copy of the probe hybridizes to the target and is cleaved by the enzyme. In a further embodiment, the process of raising and lowering the temperature is repeated many times, or cycled, as it is in PCR (Mullis and Faloona, Methods in Enzymology, 155:335 [1987], Saiki *et al.*, Science 230:1350 [1985]).

As noted above, 5' nucleases of Pol A type DNA polymerases are preferred for cleavage of an invasive cleavage structure that comprises an RNA target strand. The present invention provides enzymes having improved performance in detection assays based on the cleavage of a structure comprising RNA. In particular, the altered polymerases of the present invention exhibit improved performance in detection assays based on the cleavage of a DNA member of an invasive cleavage structure that comprises an RNA target strand.

The 5' nucleases of the present invention may be derived from Pol A type DNA polymerases. The terminology used in describing the alterations made in this class of 5' nucleases relates to the descriptions of DNA polymerase structures known in the art. The Klenow fragment of the Pol A polymerase from *E. coli* (the C-terminal two thirds, which has the DNA synthesizing activity but lacks the 5' nuclease activity) has been described as having a physical form resembling a right hand, having an open region called the "palm", and a cleft that holds the primer/template duplex defined on one side by a

5 "fingers" domain and on the other by a "thumb" domain (Joyce and Steitz, Trends in Biochemical Science 12:288 [1987]). This is shown schematically in Figure 5. Because this physical form has proved to be common to all Pol A DNA polymerases and to a number of additional template-dependent polymerizing enzymes such as reverse transcriptases, the hand terminology has become known in the art, and the sites of activity in these enzymes are often described by reference to their position on the hand. For reference, and not intended as a limitation on the present invention, the palm is created from roughly the first 200 amino acids of the polymerase domain, the thumb from the middle 140, and the fingers by the next 160, with the base of the cleft formed from the remaining regions (Figures 6). Although some enzymes may deviate from these structural descriptions, the equivalent domains and sequences corresponding to such domains may be identified by sequence homology to known enzyme sequences, by comparison of enzyme crystal structures, and other like methods.

10 In creating the improved enzymes of the present invention, several approaches have been taken, although the present invention is not limited to any particular approach. First two DNA polymerases, Taq and Tth, that have different rates of DNA strand cleavage activity on RNA targets were compared. To identify domains related to the differences in activity, a series of chimerical constructs was created and the activities were measured. This process identified two regions of the Tth polymerase that could, if transferred into the Taq polymerase, confer on the TaqPol an RNA-dependent cleavage activity equivalent to that of the native Tth protein. Once these regions were identified, the particular amino acids involved in the activity were examined. Since the two proteins are about 87 percent identical in amino acid sequence overall, the identified regions had only a small number of amino acid differences. By altering these amino acids singly and in combinations, a pair of amino acids were identified in TthPol that, if introduced into the TaqPol protein, increased the rate of cleavage up to that of the native TthPol.

25 These data demonstrate two important aspects of the present invention. First, specific amino acids can be changed to confer TthPol-like RNA-dependent cleavage activity on a polymerase having a lesser activity. More broadly, however, these results provide regions of these polymerases that are involved in the recognition of the RNA-containing cleavage structure. Identification of these important regions, combined

with published information on the relationships of other amino acids to the various functions of these DNA polymerases and computer-assisted molecular modeling during the development of the present invention have allowed a rational design approach to create additional improved 5' nucleases. The information also allowed a focused random mutagenesis approach coupled with a rapid screening procedure to quickly create and identify enzymes having improved properties. Using these methods of the present invention, a wide array of improved polymerases are provided.

The methods used in creating and selecting the improved 5' nucleases of the present invention are described in detail below and in the experimental examples. A general procedure for screening and characterizing the cleavage activity of any 5' nuclease is included in the experimental examples. The methods discussions are divided into the following sections: I) Creation and selection of chimerical constructs; II) Site-specific mutagenesis based on information from chimerical constructs; III) Site-specific mutagenesis based on molecular modeling and published physical studies; and IV) focused random mutagenesis.

1) Creation and selection of chimerical constructs

The PolA-type DNA polymerases, including but not limited to DNA polymerase enzymes from *Thermus* species, comprise two distinctive domains, the 5' nuclease and the polymerase domains, shown schematically in Figure 6. The polymerase domains reside in the C-terminal two-thirds of the proteins and are responsible for both DNA-dependent and RNA-dependent DNA polymerase activities. The N-terminal one-third portions contain the 5' nuclease domains. In the genus *Thermus* Pol A polymerase, the palm region consists of, roughly, amino acids 300-500, the thumb region includes amino acids 500-650, while the fingers region is formed by the remaining amino acids from 650 to 830 (Figure 6).

The derivatives, Taq DN RX HT and Tth DN RX HT, of Taq and TthPol used in many of the experiments of the present invention, and described herein, are modified to reduce synthetic activity and to facilitate chimera construction, but have 5' nuclease activity essentially identical to unmodified TaqPol and TthPol. Unless otherwise

specified, the TaqPol and TthPol enzymes of the following discussion refer to the DN RX HT derivative.

TthPol has a 4-fold higher cleavage rate with the IL-6 RNA template (shown in Figure 10) than TaqPol (shown in Figures 11 and 12), although the Taq and TthPols show similarities of cleavage in DNA target structures (Figure 10). Since the amino acid sequences of TaqPol and TthPol (Figures 8 and 9) share about 87% identity and greater than 92% similarity, the high degree of homology between the enzymes allowed creation of a series of chimeric enzymes between TthPol and TaqPol. The activity of the chimeric enzymes was used as a parameter to identify the region(s) of these proteins affecting RNA dependent 5' nuclease activity.

The chimeric constructs between TthPol and TaqPol genes shown schematically in Figures 7 and 19 were created by swapping DNA fragments defined by the restriction endonuclease sites, EcoRI and BamHI, common for both genes, the cloning vector site SalI and the new sites, NotI, BstBI and NdeI, created at the homologous positions of both genes by site directed mutagenesis. The restriction enzymes have been abbreviated as follows: EcoRI is E; NotI is N; BstBI is Bs; NdeI is D, BamHI is B, and SalI is S.

The activity of each chimeric enzyme was evaluated using the invasive signal amplification assay with the IL-6 RNA target (Figure 10), and the cycling cleavage rates shown in Figure 12 were determined as described in the Experimental Examples.

Comparison of the cleavage rates of the first two chimeras, TaqTth(N) and TthTaq(N), created by swapping the polymerase and 5' nuclease domains at the NotI site (Figure 7), shows that TaqTth(N) has the same activity as TthPol, whereas its counterpart TthTaq(N) retains the activity of TaqPol (Figure 12). This result indicates that the higher cleavage rate of TthPol is associated with its polymerase domain and suggests an important role of the polymerase domain in the 5' nuclease activity.

The next step was to identify a minimal region of TthPol polymerase that would give rise to the TthPol-like RNA dependent 5' nuclease activity when substituted for the corresponding region of the TaqPol sequence. To this end, the TaqTth(N) chimera was selected to generate a series of new constructs by replacing its TthPol sequence with homologous regions of TaqPol. First, the N-terminal and C-terminal parts of the TaqPol polymerase domain were substituted for the corresponding regions of TaqTth(N) using

the common BamHI site as a breaking point to create TaqTth(N-B) and TaqTth(B-S) chimeras, respectively (Figure 7). TaqTth(N-B) which has the TthPol sequence between amino acids 328 and 593, is approximately 3 times more active than the TaqTth(B-S) and 40% more active than TthPol (Figure 12). This result establishes that the NotI-BamHI portion of the TthPol polymerase domain determines superior RNA-dependent 5' nuclease activity of TthPol.

From these data it was determined that a central portion of the TthPol, when used to replace the homologous portion of TaqPol (TaqTth(N-B) construct) could confer superior RNA recognition on the chimerical protein composed primarily of Taq protein. In fact, the cycling rate of this chimerical protein exceeded that of the parent TthPol. Comparison of chimeras that included sub-portions of the activity-improving region of TthPol, approximately 50% of the region in each case (*See*, TaqTth(N-D) and TaqTth(D-B), Figures 7 and 12) showed no significant improvement in RNA dependent activity as compared to the parent TaqPol. This result indicates that aspects of each half of the region are required for this activity. A construct having an only slightly smaller portion of the Tth insert portion (TaqTth(Bs-B)) showed activity that was close to that of the parent TthPol protein, but which was less than that of the TaqTth(N-B) construct.

2) Site-specific mutagenesis based on information from chimerical constructs

Comparison of the TthPol and TaqPol amino acid sequences between the BstBI and BamHI sites reveals only 25 differences (Figure 13). Among those, there are 12 conservative changes and 13 substitutions resulting in a change in charge. Since the analysis of the chimeric enzymes has suggested that some critical amino acid changes are located in both BstBI-NdeI and NdeI-BamHI regions of TthPol, site directed mutagenesis was used to introduce the TthPol specific amino acids into the BstBI-NdeI and NdeI-BamHI regions of the TaqTth(D-B) and TaqTth (N-D) chimeras, respectively. Six TthPol-specific substitutions were generated in the BstBI-NdeI region of the TaqTth(D-B) by single or double amino acid mutagenesis and only one double mutation, W417L/G418K, was able to restore the TthPol activity with the IL-6 RNA target (*See e.g.*, Figure 14). Similarly, 12 TthPol specific amino acids were introduced at the

homologous positions of the NdeI-BamHI region of the TaqTth(N-D) and only one of them, E507Q, increased the cleavage rate to the TthPol level (*See e.g.*, Figure 14).

To confirm that the W417L, G418K and E507Q substitutions are sufficient to increase the TaqPol activity to the TthPol level, TaqPol variants carrying these mutations were created and their cleavage rates with the IL-6 RNA substrate were compared with that of TthPol. Figure 15 shows that the TaqPol W417L/G418K/E507Q and TaqPol G418K/E507Q mutants have 1.4 times higher activity than TthPol and more than 4 fold higher activity than TaqPol, whereas the TaqPol W417L/E507Q mutant has the same activity as TthPol, which is about 3 fold higher than TaqPol. These results demonstrate that K418 and Q507 of TthPol are important amino acids in defining its superior RNA dependent 5' nuclease activity compared to TaqPol.

The ability of these amino acids to improve the RNA dependent 5' nuclease activity of a DNA polymerase was tested by introducing the corresponding mutations into the polymerase A genes of two additional organisms: *Thermus filiformis* and *Thermus scotoductus*. TaqPol showed improved RNA dependent activity when it was modified to contain the W417L and E507Q mutations, which made it more similar at these residues to the corresponding residues of TthPol (K418 and Q507). The TfiPol was modified to have P420K and E507Q, creating TfiDN 2M, while the TscPol was modified to have E416K and E505Q, to create TscDN 2M. The activity of these enzymes for cleaving various DNA and RNA containing structures was determined as described in Example 1, using the IdT2, IrT3, hairpin and X-structures diagrammed in Figures 21 and 22, with the results shown in both Figure 25 and Table 8. Both enzymes have much less RNA-dependent cleavage activity than either the TthPol or the Taq 2M enzymes. However, introduction of the mutations cited above into these polymerases increased the RNA dependent cleavage activity over 2 fold compared to the unmodified enzymes (Figure 25). These results demonstrate that transferability of improved RNA dependent cleavage activity into a wide range of polymerases using the methods of the present invention.

3) Site-specific mutagenesis based on molecular modeling and published physical studies

The positions of the G418H and E507Q mutations in the crystal structure of a complex of the large fragment of TaqPol (Klentaq1) with a primer/template DNA determined by Li *et al.* (Li *et al.*, Protein Sci., 7:1116 [1998]) are shown in Figure 17. The E507Q mutation is located at the tip of the thumb subdomain at a nearest distance of 3.8 Å and 18 Å from the backbone phosphates of the primer and template strands, respectively. The interaction between the thumb and the minor groove of the DNA primer/template was previously suggested by the co-crystal structures of Klenow fragment DNA polymerase I (Breese *et al.*, Science 260:352 [1993]) and TaqPol (Eom *et al.*, Nature 382:278 [1996]). Deletion of a 24 amino acid portion of the tip of the thumb in Klenow fragment, corresponding to amino acids 494-518 of TaqPol, reduces the DNA binding affinity by more than 100-fold (Minnick *et al.*, J. Biol. Chem., 271:24954 [1996]). These observations are consistent with the hypothesis that the thumb region, which includes the E507 residue, is involved in interactions with the upstream substrate duplex.

The W417L and the G418K mutations in the palm region of TaqPol (Figure 17) are located approximately 25 Å from the nearest phosphates of the template and upstream strands, according to the co-crystal structures of TaqPol with duplex DNA bound in the polymerizing mode (Li *et al.*, Protein Sci., 7:1116 [1998], Eom *et al.*, Nature 382:278 [1996]). The same distance was observed between the analogous W513 and P514 amino acids of Klenow fragment and the template strand of DNA bound in the editing mode (Breese *et al.*, Science 260:352 [1993]). Thus, no interactions between TaqPol and the overlapping substrate can be suggested from the available co-crystal studies for this region.

Although an understanding of the mechanism of action of the enzymes is not necessary for the practice of the present invention and the present invention is not limited to any mechanism of action, it is proposed that the amino acids at positions 417 and 418 in the palm region of TaqPol interact with the upstream substrate duplex only when the enzyme functions as a 5' nuclease, but no interaction with these amino acids occurs when TaqPol switches into polymerizing mode. This hypothesis suggests a novel mode of

substrate binding by DNA polymerases called here the “5' nuclease mode.” Several lines of evidence support this hypothesis. The study of the chimeric enzymes described here clearly separates regions of the polymerase domain involved in the 5' nuclease and polymerase activities. Accordingly, the W417L and G418K mutations, together with the E507Q mutation, affect the 5' nuclease activity of TaqPol on substrates having an RNA target strand (Figure 15), but have no effect on either RNA-dependent or DNA-dependent DNA polymerase activities (Figure 16). On the other hand, mutations in the active site of TaqPol, such as R573A, R587A, E615A, R746A, N750A and D785N, which correspond to substitutions in Klenow fragment of *E.coli* DNA Pol I that affect both polymerase activity and substrate binding affinity in the polymerizing mode (Polesky *et al.*, J. Biol. Chem., 265:14579 [1990], Polesky *et al.*, J. Biol. Chem., 267:8417 [1992], Pandey *et al.*, Eur. J. Biochem., 214:59 [1993]) were shown to have little or no effect on the 5' nuclease activity. Superposition of the polymerase domains of TaqPol (Eom *et al.*, Nature 382:278 [1996]), *E.coli* Pol I and *Bacillus stearothermophilus* Pol I (Kiefer *et al.*, Nature 391:304 [1998]) using the programs DALI (Holm and Sander, J. Mol. Biol., 233:123 [1993], Holm and Sander, Science 273:595 [1996]) and Insight II (Molecular Simulation Inc., Naperville, IL) shows that the palm region of TaqPol between amino acids 402-451, including W417 and G418, is structurally highly conserved between the three polymerases, although there is no structural similarity between the rest of the palm subdomains. This observation suggests an important role for this region in eubacterial DNA polymerases.

The 5' nuclease and polymerase activities should be precisely synchronized to create a nicked structure rather than a gap or an overhang that could cause a deletion or an insertion during Okazaki fragment processing or DNA repair, if ligase joins the ends inappropriately. According to the previously proposed model (Kaiser *et al.*, J. Biol. Chem., 274:21387 [1999]), the 3' terminal nucleotide of the upstream strand is sequestered by the 5' nuclease domain to prevent its extension, thus halting synthesis. The interaction with the 3' nucleotide apparently activates the 5' nuclease that endonucleolitically removes the displaced 5' arm of the downstream strand. This cleavage occurs by the precise incision at the site defined by the 3' nucleotide, thus creating the nick. This model requires a substantial rearrangement of the

substrate-enzyme complex, which may include a translocation of the complex to the 5' nuclease mode to separate the primer/template from the polymerase active site.

It is possible that a relocation of the substrate away from the polymerase active site could be induced by the interaction between the duplex formed between the template and incoming strands and the crevice formed by the finger and thumb subdomains. Such an interaction could force conformational transitions in the thumb that would bring the template/primer duplex into close contact with the W417 and G418 amino acids.

Significant flexibility of the thumb has been previously reported that might explain such changes (Beese *et al.*, Science 260:352 [1993], Eom *et al.*, Nature 382:278 [1996], Ollis *et al.*, Nature 313:762 [1985], Kim *et al.*, Nature 376:612 [1995], Korolev *et al.*, Proc. Natl. Acad. Sci., 92:9264 [1995], Li *et al.*, EMBO J., 17:7514 [1998]). Additional conformational changes of the fingers domain that might help to open the crevice, such as the transition from the 'closed' to the 'open' structure described by Li *et al.* (Li *et al.*, EMBO J., 17:7514 [1998]), are consistent with this model. It may be that the 5' nuclease binding mode was not observed in any of the published co-crystal structures of a DNA Pol I because the majority of the structures were solved for the polymerase domain only, with a template/primer substrate rather than with an overlapping 5' nuclease substrate.

K_m values of 200-300 nM have been determined for TaqPol, TthPol and TaqPol G418K/E507Q for the RNA containing substrate. These values are much higher than the K_m value of <1 nM estimated for TthPol with an all-DNA overlapping substrate suggesting that the RNA template adversely affects substrate binding. The low affinity could be explained by the unfavorable interaction between the enzyme and either the A-form duplex adopted by the substrate with an RNA target, or the ribose 2' hydroxyls of the RNA strand. Between these two factors, the latter seems more likely, since the 5' nucleases of eubacterial DNA polymerases can efficiently cleave substrates with an RNA downstream probe (Lyamichev *et al.*, Science 260:778 [1993]), which would presumably have an A-form. Further, the co-crystal studies suggest that the template/primer duplex partially adopts a conformation close to A-form in its complex with DNA polymerase (Eom *et al.*, Nature 382:278 [1996], Kiefer *et al.*, Nature 391:304 [1998], Li *et al.*, EMBO J., 17:7514 [1998]). The G418K/E507Q mutations increase the k_{cat} of TaqPol more than two fold, but have little effect on K_m . Such an effect would be expected if the

mutations position the substrate in an orientation more appropriate for cleavage rather than simply increasing the binding constant.

In addition to the mutational analysis described above, another approach to studying specific regions of enzymes, enzyme structure-function relationships, and enzyme-substrate interaction is to investigate the actual, physical structure of the molecule.

With the advances in crystallographic, NMR, and computer and software technology, study of molecular structure has become a viable tool for those interested in the configuration, organization, and dynamics of biomolecules. Molecular modeling has increased the understanding of the nature of the interactions that underlie the structure of proteins and how proteins interact functionally with substrate. Numerous publications describing the structures of various polymerases or polymerase protein portions, HIV reverse transcriptase, and other nucleic acid binding proteins have provided mechanistic insights into protein conformation, changes in conformation, and molecular interactions necessary for function.

As an example, the report by Doublie *et al.* (Doublie *et al.*, Nature 391:251 [1998]) discloses the crystal structure of T7 DNA polymerase and provides information about which amino acid regions are likely to have an affect on substrate binding, which are required to contact the substrate for polymerization, and which amino acids bind cofactors, such as metal ions. It is noted in this paper and others that many of the polymerases share not only sequence similarity, but structural homology as well. When certain structural domains of different polymerases are superimposed (for example, T7 polymerase, Klenow fragment editing complex, the unliganded Taq DNA polymerase and the Taq Polymerase-DNA complex) conserved motifs are clearly discernable.

Specifically, combining the information from all of these different structural sources and references, a model of the protein interacting with DNA, RNA, or heteroduplex can be made. The model can then be examined to identify amino acids that may be involved in substrate recognition or substrate contact. Changes in amino acids can be made based on these observations, and the effects on the various activities of the 5' nuclease proteins are assessed using screening methods such as those of the present invention, described in the experimental examples.

The domain swapping analysis discussed previously demonstrated that sequences of TthDN that are important in RNA-dependent 5' nuclease activity lie in the polymerase domain of the protein. Therefore, study of structural data of the polymerase domain with respect to nucleic acid recognition provides one method of locating amino acids that,

- 5 when altered, alter RNA recognition in a 5' nuclease reaction. For example, analysis conducted during the development of the present invention examined published analyses relating to primer/template binding by the polymerase domain of *E. coli* Pol I, the Klenow fragment. Table 2 shows a sampling of kinetic constants determined for the Klenow fragment, and shows the effects a number of mutations on these measurements.
- 10 The corresponding or similarly positioned amino acids in the TaqPol are indicated in the right hand column. It was postulated that mutations having a noticeable impact on the interactions of the Klenow fragment with the DNA template or the primer/template duplex, as indicated by changes in K_d and Relative DNA affinity values, might also have effects when made at the corresponding sites in TaqPol and related chimerical or mutant
- 15 derivatives. A selection of the mutations that produced a higher K_d value or a lower Relative DNA affinity value when introduced into the Klenow fragment were created and examined in TaqPol. These Taq derivatives include, but are not limited to, those indicated by asterisks in the right hand column of Table 2.

- For some Klenow variants, such as the R682 mutants, selection for testing was
- 20 not made based on the DNA affinity measurements, but because molecular modeling suggested interaction between some aspect of the template/primer duplex and that amino acid. Similarly, additional regions of Taq polymerase (or Taq derivatives) were targeted for mutagenesis based on structural data and information from molecular modeling.
- Based on modeling, the thumb region was postulated to contact an RNA template. Thus,
- 25 amino acids in the thumb region were looked for that, if altered, might alter that contact. For example, Figures 6 and 17 show that amino acids 502, 504, and 507 are located at the tip of the thumb. It was postulated that altering these amino acids might have an affect on the enzyme-substrate interaction. Using the activity screening methods described In Example 1, mutations that produced beneficial effects were identified. This approach
- 30 was used to create a number of improved enzymes. For example, TaqPol position H784, corresponding to Klenow amino acid H881, is an amino acid in the fingers region and, as

such, may be involved in primer/template substrate binding. When the H881 amino acid in the Klenow enzyme is replaced by alanine, the change decreases the affinity of the enzyme for DNA to only 30 to 40% of the wild type level. An analogous substitution was tested in a TaqPol-derived enzyme. Starting with the Taq derivative

5 W417L/G418K/E507Q, amino acid 784 was changed from Histidine (H) to Alanine (A) to yield the W417L/G418K/E507Q/H784A mutant, termed Taq 4M. This variant showed improved 5' nuclease activity on the RNA test IrT1 (Figure 24) test substrate (data in Table 3). Amino acid R587 is in the thumb region, and was selected for mutation based on its close proximity to the primer/template duplex in computer models. When an
10 R587A mutation was added to the Taq 4M variant, the activity on the test IrT1 test substrate was still further improved. In addition, the reduction, relative to the 4M variant, in cleavage of the X structure shown in Figure 22 constitutes an additional improvement in this enzyme's function.

Not all amino acid changes that reduce DNA binding in the polymerization affect
15 the 5' nuclease activity. For example, mutations E615A, R677A, affecting amino acid that are also in the thumb and fingers domains, respectively, have either adverse effect, or no effect on the 5' nuclease activities, respectively, as measured using the test substrates in Figures 21 and 22, and compared to the parent variants that lacked these changes. The R677A mutation was added to, and compared with the TaqSS variant, while the E615A
20 mutation was added to and compared with the Taq 4M variant. The test methods described herein provide a convenient means of analyzing any variant for the alterations in the cleavage activity of both invasive and noninvasive substrates, for both DNA and RNA containing structures. Thus, the present invention provides methods for identifying all suitable improved enzymes.

25 Alterations that might increase the affinity of the enzymes for the nucleic acid targets were also examined. Many of the mutations described above were selected because they caused the Klenow fragment enzyme to have decreased affinity for DNA, with the goal of creating enzymes more accepting of structures containing non-DNA strands. In general, the native DNA polymerases show a lower affinity for RNA/DNA
30 duplexes, compared to their affinity for DNA/DNA duplexes. During the development of the present invention, it was sought to increase the general affinity of the proteins of the

present invention for a nucleic acid substrate without restoring or increasing any preference for structures having DNA rather than RNA target strands. The substitution of amino acids having different charges was examined as a means of altering the interaction between the proteins and the nucleic acid substrates. For example, it was postulated that addition of positively charged amino acid residues, such as lysine (K), might increase the affinity of a protein for a negatively charged nucleic acid.

As noted above, alterations in the thumb region could affect the interactions of the protein with the nucleic acid substrate. In one example, the mutation G504K (tip of the thumb) was introduced in Taq4M and caused an enhancement of nuclease activity by 15% on an RNA target. Additional positively charged mutations (A502K and E507K) further improve the RNA target dependent activity by 50% compared to the parent Taq4M enzyme.

The use of data from published studies and molecular modeling, in combination with results accrued during the development of the present invention allowed the identification of regions of the proteins in which changes of amino acids would be likely to cause observable differences in at least one aspect of cleavage function. While regions could be targeted in this way, it was observed that changes in different amino acids, even if near or immediate neighbors in the protein, could have different effects. For example, while the A502K substitution created a marked increase in the RNA-dependent cleavage activity of Taq 4M, changing amino acid 499 from G to a K, only 3 amino acids away from 502, gave minimal improvement. As can be seen in the Experimental Examples, the approach of the present invention was to change several amino acids in a candidate region, either alone or in combination, then use the screening method provided in Example 1 to rapidly assess the effects of the changes. In this way, the rational design approach is easily applied to the task of protein engineering.

In addition to the thumb, palm, and hand regions found in the polymerase domain of these proteins, regions that are specific to 5' nucleases and nuclease domains were examined. Comparative studies on a variety of 5' nucleases have shown that, though the amino acid sequences vary dramatically from enzyme to enzyme, there are structural features common to most. Two of these features are the helix-hairpin-helix motif (H-h-H) and the arch or loop structure. The H-h-H motif is believed to mediate

non-sequence specific DNA binding. It has been found in at least 14 families of proteins, including nucleases, N-glycosylases, ligases, helicases, topoisomerases, and polymerases (Doherty *et al.*, Nucl. Acid. Res., 24:2488 [1996]). The crystallographic structure of rat DNA polymerase pol β bound to a DNA template-primer shows non-specific hydrogen
5 bonds between the backbone nitrogens of the pol β HhH motif and the phosphate oxygens of the primer of the DNA duplex (Pelletier *et al.*, Science 264:1891 [1994]). Because the HhH domain of 5' nuclease domains of Taq and Tth polymerases may function in a similar manner, it is contemplated that mutations in the HhH region of the enzyme alter activity. Mutations may be introduced to alter the shape and structure of the
10 motif, or to change the charge of the motif to cause increased or decreased affinity for substrate.

Another structure common to many 5' nucleases from diverse sources such as eukaryotes, eubacteria, archaea and phage, is the arch or loop domain. The crystal structure of the 5' exonuclease of bacteriophage T5 showed a distinct arch formed by two
15 helices, one positively charged and one containing hydrophobic residues (Ceska *et al.*, Nature 382:90 [1996]). Interestingly, three residues that are conserved between T5 and Taq, Lys 83, Arg 86 and Tyr 82 are all in the arch. These correspond to amino acids Lys 83, Arg 86, and Tyr 82 in Taq DNA polymerase. The crystal structure for Taq (5' nuclease) has also been determined (Kim *et al.*, Nature 376:612 [1995]).

The crystal structure from the flap endonuclease-1 from *Methanococcus janneschii* also shows such a loop motif (Hwang *et al.*, Nat. Struct. Biol., 5:707 [1998]). The backbone crystal structure of Mja FEN-1 molecules may be superimposed on T5 exonuclease, Taq 5'-exonuclease and T4 RNase H. An interesting feature common to all of these is the long loop. The loop of FEN-1 consists of a number of positively charged
25 and aromatic residues and forms a hole with dimensions large enough to accommodate a single-stranded DNA molecule. The corresponding region in T5 exonuclease consists of three helices forming a helical arch. The size of the hole formed by the helical arch in T5 exonuclease is less than half that formed by the L1 loop in Mj FEN-1. In T4 RNase H or Taq 5' exonuclease, this region is disordered. Some regions of the arch bind metals,
30 while other regions of the arch contact nucleic acid substrate. Alignment of the amino-acid sequences of six 5' nuclease domains from DNA polymerases in the pol I

family show six highly conserved sequence motifs containing ten conserved acidic residues (Kim *et al.*, Nature 376 [1995]).

The effects of alterations in the arch region were examined. In Taq polymerase the arch region is formed by amino acids 80-95 and 96-109. Site directed mutagenesis was performed on the arch region. Alignment of amino acid sequences of the FEN and polymerase 5' nucleases suggested the design of 3 amino acid substitution mutations, P88E, P90E and G80E. These substitutions were made on the Taq4M polymerase mutant as a parent enzyme. Results indicated that although the background activity on the HP and X substrates shown in Figure 22 are tremendously suppressed in all mutants, the desirable 5' nuclease activity on proper substrates (IdT and IrT, Figure 24) is also reduced. Despite the sequence homology between Taq and Tth polymerases, they have very different activity on HP and X substrates. The alignment of the Taq and Tth polymerase arch regions also demonstrates regions of extensive sequence homology as well as minor differences. These differences led to the design of mutations L109F and A110T using Taq4M to generate Taq4M L109F/A110T, and the mutant Taq 4M A502K/G504K/E507K/T514S to generate Taq 4M L109F/A110T/A502K/G504K/E507K/T514S mutant. These two mutations have drastically converted Taq4M enzyme to become more like Tth enzyme in terms of the background substrate specificity while the 5' nuclease activities on both DNA and RNA substrates are almost unchanged.

4) Focused random mutagenesis

As described above, physical studies and molecular modeling may be used alone or in combination to identify regions of the enzymes in which changes of amino acids are likely to cause observable differences in at least one aspect of cleavage function. In the section above, use of this information was described to select and change specific amino acids or combinations of amino acids. Another method of generating an enzyme with altered function is to introduce mutations randomly. Such mutations can be introduced by a number of methods known in the art, including but not limited to, PCR amplification under conditions that favor nucleotide misincorporation (REF), amplification using primers having regions of degeneracy (*i.e.*, base positions in which different individual,

but otherwise similar oligonucleotides in a reaction may have different bases), and chemical synthesis. Many methods of random mutagenesis are known in the art (Del Rio *et al.*, *Biotechniques* 17:1132 [1994]), and may be incorporated into the production of the enzymes of the present invention. The discussions of any particular means of

5 mutagenesis contained herein are presented solely by way of example and not intended as a limitation. When random mutagenesis is performed such that only a particular region of an entire protein is varied, it can be described as "focused random mutagenesis." As described in the Experimental Examples, a focused random mutagenesis approach was applied to vary the HhH and the thumb domains some of the enzyme variants previously
10 created. These domains were chosen to provide examples of this approach, and it is not intended that the random mutagenesis approach be limited to any particular domain, or to a single domain. It may be applied to any domain, or to any entire protein. Proteins thus modified were tested for cleavage activity in the screening reactions described in Example 1, using the test substrates diagrammed in Figures 22 and 24, with the results
15 described in Tables 6 and 7.

Random mutagenesis was performed on the HhH region with the parent TaqSS or TthDN H785A mutants. None of the 8 mutants generated showed an improvement in activity compared to the parent enzyme (Table 6). In fact, mutations of the region between residues 198-205 have about 2-5 fold lower activity on both DNA and RNA
20 substrates, suggesting that this region is essential for substrate recognition. Mutagenesis in the thumb region resulted in new mutations that improved 5' nuclease activity by 20-100% on a DNA target and about 10% on an RNA target (Table 7).

Numerous amino acids in each of the distinct subdomains play roles in substrate contact. Mutagenesis of these may alter substrate specificity by altering substrate
25 binding. Moreover, mutations introduced in amino acids that do not directly contact the substrate may also alter substrate specificity through longer range or general conformation altering effects. These mutations may be introduced by any of several methods known in the art, including, but not limited to random mutagenesis, site directed mutagenesis, and generation of chimeric proteins.

30 As noted above, numerous methods of random mutagenesis are known in the art. The methods applied in the focused random mutagenesis described herein may be applied

to whole genes. It is also contemplated that additional useful chimerical constructs may be created through the use of molecular breeding (*See e.g.*, U.S. Pat. No. 5,837,458 and PCT Publications WO 00/18906, WO 99/65927, WO 98/31837, and WO 98/27230, herein incorporated by reference in their entireties). Regardless of the mutagenesis method chosen, the rapid screening method described herein provides a fast and effective means of identifying beneficial changes within a large collection of recombinant molecules. This makes the random mutagenesis procedure a manageable and practical tool for creating a large collection of altered 5' nucleases having beneficial improvements. The cloning and mutagenesis strategies employed for the enzymes used as examples are applicable to other thermostable and non-thermostable Type A polymerases, since DNA sequence similarity among these enzymes is very high. Those skilled in the art would understand that differences in sequence would necessitate differences in cloning strategies, for example, the use of different restriction endonucleases may be required to generate chimeras. Selection of existing alternative sites, or introduction via mutagenesis of alternative sites are well established processes and are known to one skilled in the art.

Enzyme expression and purification can be accomplished by a variety of molecular biology methods. The examples described below teach one such method, though it is to be understood that the present invention is not to be limited by the method of cloning, protein expression, or purification. The present invention contemplates that the nucleic acid construct be capable of expression in a suitable host. Numerous methods are available for attaching various promoters and 3' sequences to a gene structure to achieve efficient expression.

5) Site-Specific mutagenesis

In some embodiments of the present invention, any suitable technique (*e.g.*, including, but not limited to, one or more of the techniques described above) are used to generate improved cleavage enzymes (*e.g.*, SEQ ID NO:221) with heterologous domains. Accordingly, in some embodiments, site-specific mutagenesis (*e.g.*, primer-directed mutagenesis using a commercially available kit such as the Transformer Site Directed mutagenesis kit (Clontech)) is used to make a plurality of changes throughout a nucleic acid sequence in order to generate nucleic acid encoding a cleavage enzyme of the

present invention. In some embodiments, a plurality of primer-directed mutagenesis steps are carried out in tandem to produce a nucleic acid encoding a cleavage enzyme of the present invention.

In some embodiments, a plurality of primer directed mutagenesis steps are directed to a selected portion of a nucleic acid sequence, to produce changes in a selected portion of a cleavage enzyme of the present invention. In other embodiments, a nucleic acid having changes in one selected portion is recombined with a nucleic acid having mutations in a different selected portion (*e.g.*, through cloning, molecular breeding, or any of the other recombination methods known in the art), thereby creating a nucleic acid having mutations in a plurality of selected portions, and encoding a cleavage enzyme of the present invention. The mutations in each selected portion may be introduced by any of the methods described above, or any combination of said methods, including but not limited to methods of random mutagenesis and site-directed mutagenesis.

For example, in one illustrative embodiment of the present invention, the nucleic acid sequence of SEQ ID NO:222 (a nucleic acid sequence encoding the cleavage enzyme of SEQ ID NO:221) is generated by making a plurality of primer-directed mutations to the nucleic acid sequence of SEQ ID NO: 104 (see Example 7 for the construction of SEQ ID NO:104). In some embodiments, each mutation is introduced using a separate mutagenesis reaction. Reactions are carried out sequentially such that the resulting nucleic acid (SEQ ID NO: 222) contains all of the mutations. In another illustrative embodiment of the present invention, the nucleic acid sequence of SEQ ID NO:222 is generated by making a plurality of primer-directed mutations, as described above, in the nuclease portion (*e.g.*, as diagrammed in Figure 6) of SEQ ID NO:111. The mutant nuclease portion is then combined with the "polymerase" portion of SEQ ID NO:104 at the Not I site, using the recombination methods described in Example 4, thereby creating a single nucleic acid having SEQ ID NO:222, and encoding the cleavage enzyme of SEQ ID NO:221. Following mutagenesis, the resulting altered polypeptide is produced and tested for cleavage activity using any suitable assay (*e.g.*, including, but not limited to, those described in Examples 1 and 6). In some embodiments, the nucleic acid sequence encoding the cleavage enzyme of SEQ ID NO:221 (*e.g.*, SEQ ID NO:222) is further modified using any suitable method.

V. Reaction Design for INVADER Assay Detection of RNA Targets;

Approaches to designing INVADER assays for the detection of RNA targets can vary depending on the needs of a particular assay. For example, in some embodiments, an RNA to be detected or analyzed may be present in a test sample at low levels, so a high level of sensitivity (*i.e.*, a low limit of detection, or LOD) may be desirable; in other embodiments, an RNA may be abundant, and may not require an especially sensitive assay for detection. In some embodiments, an RNA to be detected may be similar to other RNAs in a sample that are not intended to be detected, so that a high level of selectivity in an assay is desirable, while in other embodiments, it may be desired that multiple similar RNAs be detected in a single reaction, so an assay may be provided that is not selective with respect to the differences among these similar RNAs.

In some embodiments it is especially desirable to avoid detection of any DNA molecules related to the target RNA molecules in a reaction. In some embodiments, this is accomplished by designing INVADER assay probe sets to RNA splice junctions, such that only the properly spliced mRNAs provide the selected target sites for detection. In other embodiments, samples are handled such that DNA remains double stranded (*e.g.*, the nucleic acids are not heated or otherwise subjected to denaturing conditions), and is thus not available to serve as target in an INVADER assay reaction. In other embodiments, cells are lysed under conditions that leave nuclei intact, thereby containing and preventing detection of the genomic DNA, while releasing the cytosolic mRNAs into the lysate solution for detection by the assay.

In some embodiments, the INVADER assay is to be used for detection or quantitation of an entire RNA having a particular variation of a sequence (*e.g.*, a mutation a SNP, a particular spliced junction); in such embodiments, the location of the base or sequence to be detected is a determining factor in the selection of a site for the INVADER assay probe set to hybridize. In other embodiments, any portion of an RNA target may be used to indicate the presence or the amount of the entire RNA (*e.g.*, as in gene expression analysis). In this case, the probe sets may be directed toward a portion of the RNA selected for optimal performance (*e.g.*, sites determined to be particularly accessible for probe hybridization) as a target in the INVADER assay.

The discussion of INVADER assay probe design is divided into the following sections:

- i. Target site selection based on accessibility
- ii. Target site selection based on selectivity
- 5 iii. Oligonucleotide design
 - a. Target-specific regions: length and melting temperature
 - b. Non-complementary regions
 - c. Folding and dimer analysis
- iv. Assay performance evaluation
- 10 v. Design and assay optimization

- i. **Target site selection based on accessibility**

One consideration in the selection of sites for detection is the availability of the target site for hybridization of the assay probe set. To simply use randomly selected complementary oligonucleotides for a given RNA target without prior knowledge of regions of the RNA that allow efficient hybridization can be an ineffective approach. For example, it is estimated that targeting RNA with antisense oligonucleotides based on random design results in one out of 18-20 tested oligonucleotides showing significant inhibition of gene expression (Sczakiel, *Frontiers in Biosciences* 5:194 [2000]; Patzel *et al.*, *Nucleic Acids Res.*, 27:4328 [1999]; Peyman *et al.*, *Biol. Chem. Hoppe-Seyler* 367:195 [1995]; Monia *et al.*, *Nature Med.*, 2:668 [1996]). Secondary and tertiary structures of RNA are thought to be the major reasons that influence the ability of an oligonucleotide to bind targeted regions of the RNA (Vickers *et al.*, *Nucleic Acids Res.*, 28:1340 [2000]; Lima *et al.*, *Biochemistry* 31:12055 [1992]; Uhlenbeck, *J. Mol. Biol.*, 65:25 [1972]; Freier and Tinoco, *Biochemistry* 14:3310 [1975]). This is due to the hybridization kinetics and thermodynamics of destroying any structural motifs of the RNA and, in return, hybridizing the complementary DNA oligonucleotide (Patzel *et al.*, *Nucleic Acids Res.*, 27:4328 [1999]; Mathews *et al.*, *RNA* 5:1458 [1999]). Thus, the ability to identify regions of RNA that are “accessible” for hybridization is important for design and selection of effective oligonucleotides.

There are several experimental and theoretical methods available for identifying accessible regions in RNA. These include the use of RNase-H footprinting (Ho *et al.*, Nature Biotechnology 16:59 [1998]; Mateeva *et al.*, Nucleic Acids Res., 25:5010 [1997]; Mateeva *et al.*, Nature Biotechnology 16:1374 [1998]), complementary arrays of

5 oligonucleotide libraries (Southern *et al.*, Nucleic Acids Res., 22:1368 [1994]; Mir and Southern, Nature Biotechnology 17:788 [1999]), ribozyme libraries with random hexamer internal guide sequences (Campbell and Cech, RNA 1:598 [1995]; Lan *et al.*, Science 280:1593 [1998]), and RNA and DNA structure prediction computer programs (Sczakiel, Frontiers in Biosciences 5:194 [2000]; Patzel *et al.*, Nucleic Acids Res.,

10 27:4328 [1999]; Zuker, Science 244:48 [1989]; Walton *et al.*, Biotechnol. Bioeng., 65:1 [1999]). Recently, new methods have been developed that use primer extension to identify sites in RNAs that are accessible for hybridization. Target nucleic acids (*e.g.*, mRNA target nucleic acids) are contacted with a plurality of primers containing a 3' a region of degenerate sequence and primer extension reactions are conducted. Where the

15 target nucleic acid is an RNA molecule, preferred enzymes for use in the extension reactions are reverse transcriptases, which produce a DNA copy of the RNA template. Folded structures present in the target nucleic acid affect the initiation and/or efficiency of the extension reaction. The extension products of the primers are analyzed to provide a map of the accessible sites. For example, certain extension products are not generated

20 where the primer is complementary to a sequence that is involved in a folded structure. Regions of the target nucleic acid that do not allow hybridization of the primer and do not result in the production of an extension product are considered inaccessible sites. In contrast, the presence of an extension product indicates that the primer was able to bind to an accessible region of the target nucleic acid. Such methods are referred to herein as

25 "reverse transcription with random oligonucleotide libraries" or "RT-ROL" (HT Allawi, *et al.*, RNA 7(2):314-27 [2001]). The use of a physical measurement such as RT-ROL or array hybridization provides the most direct evidence of the accessibility of a site on an RNA strand. In general, INVADER assays directed toward accessible regions produce stronger signals for a given amount of RNA than assays directed toward less accessible

30 regions of an RNA strand. For the detection of rare RNAs (*e.g.*, fewer than about 5,000 to 10,000 copies per INVADER assay reaction), or in any assay wherein it is desirable to

have the best (*i.e.*, lowest) limit of detection possible, it may be beneficial to start the assay design by analyzing the RNA structure using RT-ROL or another method of physical analysis.

In other embodiments, ease of assay design may be more important than creation of an assay with a particularly low LOD. Structure prediction software can simplify the task of determining which parts of an RNA are likely to be single stranded, and thus be more accessible for probe hybridization. As first step, the sequence of an RNA to be detected is entered into an electronic file. It may be entered manually or imported from a file (*e.g.*, a sequence data file, or a word processing file). In some embodiments, the sequence is downloaded from a database, such as GenBank or EMBL. The RNA sequence can then be analyzed using a program such as *mfold* (Zucker, Science 244:48 [1989]), OligoWalk (Mathews *et al.*, RNA 5:1458 [1999]), and variations of both (Sczakiel, Frontiers in Biosciences 5:194 [2000]; Patzel *et al.*, Nucleic Acids Res., 27:4328 [1999]; Walton *et al.*, Biotechnol. Bioeng., 65:1 [1999]).

***Mfold* Analysis for target RNA structure prediction.**

The output of *mfold* analysis can be used in several ways to assist in identifying accessible regions of an RNA target molecule. In one embodiment, the *mfold* program is used to generate an "ss count" file for identifying regions least likely to be involved in intra-strand basepairing. In another embodiment, the *mfold* program is used to generate a ".ct" file, a file used as input information for use with RNA Structure 3.5 to perform an OligoWalk analysis. In preferred embodiments, for either use, the sequence to be detected is entered into *mfold*. In a preferred embodiment, the settings used in the *mfold* analysis include:

- Folding Temperature: 37°C (Even though the INVADER reaction may not be conducted at this temperature.)
- % Suboptimality: 5
- # foldings: 50
- Window Parameter: Default
- Maximum distance between paired bases: No Limit
- Select BATCH folding

- Enter: an e mail address where the results are to be sent when ready
- Image Resolution: High
- Structure Format: Bases
- Base Number Frequency: Default
- 5 • Structure Rotation Angle: 0
- Structure Annotation: SS-Count
- 1M NaCl (Australian *mfold* Internet site only)

When results are ready, an e mail message is sent containing the Web address of the results. The only file that is necessary for subsequent INVADER assay probe design analysis is the SS-Count file, which is then downloaded from the Results page. An exemplary *mfold* analysis using a GenBank entry for Human Ubiquitin (#4506712) is shown below:

SS-Count analysis for accessible sites identification.

The SS-Count file is then imported into an Excel spreadsheet file and the following options are chosen: Data Type = Delimited <press Next>; Delimited = Select (x) Spaces; (x) Treat multiple delimiters as one <press Next>; Column for data format = General. Selecting these options results in the import into Excel of three columns of data (Figure 39). The first line in the first column is the total number of stable structures *mfold* found under the parameters used in the folding. With the Ubiquitin example, there were 12 structures found.

The rest of the first column is the RNA nucleotide position numbered from 1. The second column is the raw SS-Count number and represents the number of times the corresponding base was NOT base-paired as part of some secondary structure. The third column is the sequence of the RNA analyzed, identifying the base at each numbered position. By looking for bases that are involved in fewer structures (*i.e.*, bases listed in column 3 corresponding to the higher numbers in column 2), it is possible to identify regions of the mRNA (identified by position number in column 1) that are more likely to be free of intra-strand base-pairing, and thus are more likely to be available for detection using the INVADER assay.

One way of viewing the data is to calculate the running average SS-count for a ten nucleotide stretch of the RNA (the Ave(10) Index) and chart the Ave(10) Index against the base pair position (See e.g., Figure 48).

An alternative plot is to graph the nucleotide position against the Ave(10) Index expressed as a percentage of the total number of structures found by *mfold* (Fig 39). As with the raw SS-count table, regions of the RNA corresponding to higher numbers in the Ave(10) Index are involved in fewer predicted folded structures. Viewing the Ave(10) Index in either a chart or graph format reduces the complexity of the data, and can reveal longer stretches of the RNA that are more likely to be structure-free. For example, from the graph of the running average, a user can pick out all of the major peak areas as likely regions for INVADER assay probe design. This creates an SS-Count Candidate List. In the Human ubiquitin example, there is one major peak and about 6 other peaks (Fig. 39)

The next step is to refer back to the raw (*i.e.*, not the running average) SS-Count data from within each of the peak areas, and identify the residue where the running average is changing in magnitude, this is a local "turn" and is generally a good candidate residue to be positioned at the INVADER assay probe set cleavage site. For example, in Human Ubiquitin, residue G119 is found at a minor local turn within a globally accessible region (Fig. 40). The INVADER assay probe set with the cleavage site at this location is a good performer in detection of this RNA. Placing the cleavage site at G114 did not result in better detection even though it had a higher %Ave(10) value. While not limiting the present invention to any particular mechanism, and an understanding of the mechanism is not necessary to practice the invention, this is likely to be because this area was less accessible to the Probe or INVADER oligonucleotides than was position G119.

OligoWalk Structure prediction with RNA Structure 3.5 for accessible sites

identification

In some embodiments, the program OligoWalk, a module of the software "RNAstructure" (Mathews *et al.*, RNA 5:1458 [1999]) is used in the selection of sites that are more likely to be accessible for oligonucleotide binding. OligoWalk uses sets of thermodynamic parameters for both RNA and DNA, and their hybrids (Allawi and SantaLucia, Biochemistry 36:10581 [1997]; Mathews *et al.*, J. Mol. Biol., 288:911

[1999]; Sugimoto *et al.*, Biochemistry 34:11211 [1995]) in an algorithm that relies on *mfold* for RNA secondary structure prediction (Zucker, Science 244:48 [1989]).

OligoWalk is designed to predict the most favorable regions of an RNA target for designing antisense oligonucleotides by estimating the overall thermodynamics of

5 hybridizing an antisense oligomer to the RNA by taking into account the thermodynamics of destroying any structural motifs in the RNA target or the antisense oligonucleotide.

The affinity of the oligomer to its target is expressed as an overall Gibbs free energy change of a self-structured oligomer, and of a target associating into an oligomer–target complex. This free energy is usually a negative number, indicating favorable binding,

10 and is expressed in 'kcal/mol' units. OligoWalk analysis is performed with 8 to 15 base oligonucleotide size to resemble the average length of the analyte specific region of the Signal Probe. Plotting the total binding energy against the length of the RNA generates a graph of peaks and valleys. The lowest negative values generally indicate the most favorable sites for oligonucleotides to bind. The most inaccessible regions have positive
15 binding energy values, and generally are a poor sites for assay probe design

In a preferred embodiment, the OligoWalk module of RNA Structure 3.5 is used to determine binding energies by performing an 8-base OligoWalk using the following settings:

- Break Local Structure
- 20 ▪ Include suboptimal structures
- Oligo Length: 8nt
- Oligo Concentration: 100 nM
- Oligo Type: DNA
- Walk entire Target RNA

25

When these parameters have been set, the sequence file to be folded (the ".ct" output file from *mfold*) can be selected and opened. Once the sequence has been folded, a report can be created using the Output menu. The report is imported into Excel and the data generated above is plotted. In a preferred embodiment, the OligoWalk data is
30 graphed with the SS-Count data. The regions displaying the lowest free energy values (*i.e.*, the largest negative numbers) are generally the most likely to be accessible for

hybridization. In preferred embodiments, the 3' end and the majority of the target-binding region of the probe oligonucleotide complement an accessible region of the target RNA. In particularly preferred embodiments, the majority of the binding site for the corresponding INVADER oligonucleotide falls within the same accessible region. In
5 another preferred embodiment, the binding site for an INVADER oligonucleotide falls within a nearby accessible region.

An INVADER oligonucleotide can generally be positioned to bind to a less accessible site. While not limiting the present invention to any particular mechanism, it is observed that the INVADER oligonucleotides are generally longer than probe
10 oligonucleotides used in the INVADER assay reactions and, because they are generally designed to remain bound to the target at the reaction temperature, they will be selected to have a T_m s about 12 to 15 °C higher than that of a corresponding probe. Consequently, INVADER oligonucleotides may more readily break the local target structure, and thus may be less dependent on the accessibility of the target-binding site.

15 In selecting among accessible sites for the design of INVADER assay oligonucleotides, the base composition of the site is also considered. It has been observed that stretched of more than 4 or 5 of the same nucleotide in a row (*e.g.*, ...AAAA... or ...CCCC...) in any portion of the binding site for the assay oligonucleotides may reduce the performance of the probe set in the assay (*e.g.*, by
20 increasing background or decreasing specificity). Thus, in preferred embodiments, any stretches comprising four or more repeated bases are generally avoided. Another consideration is the effect of base composition on lengths of the oligonucleotides in the probe set. In many cases, targeting A-T rich sequences requires the use of longer oligonucleotides for a reaction performed at a given temperature, compared to the length
25 of oligonucleotides targeted to sequences having a more even distribution of A-T and G-C bases. Longer oligonucleotides can be more prone to formation of intrastrand structures and dimer structures. Thus, it is preferred that the distribution of A-T bases and G-C bases within a target region be as close to even (*i.e.*, about 50% G-C content) as the region to be detected permits. In particularly preferred embodiments, the distribution
30 of A-T and G-C positions is evenly distributed across the binding sites (*e.g.*, not having all A-T positions in one half, with all G-C positions in the other).

ii. Target site selection based on selectivity

In some embodiments, probe sets are designed to examine highly homologous, or closely related RNA targets (*i.e.*, targets that are very similar in sequence). In such
5 embodiments, the RNA or homologous cDNA sequences are compared, *e.g.*, using an alignment program such as MEGALIGN (DNASTar Madison, WI).

In some embodiments, selectivity is provided by designing probe sets to detect splice junctions. Splice junctions can be identified by aligning the cDNA and gene sequences using an alignment program (*e.g.* MEGALIGN) or under the BLAST menu at
10 the NCBI website (BLAST 2 sequences). Splice junctions are also often listed in the GenBank report (intron/exon sites). INVADER assay oligonucleotide sets are designed such that the probe and INVADER oligonucleotides are complementary to the coding strand (mRNA), generally with the cleavage site being as close to the splice junction as possible. In some embodiments, different splice junctions within an mRNA are analyzed
15 for accessibility, as described above. In preferred embodiments, probe sets are designed to detect one or more splice junctions showing greater accessibility compared to the accessibility of other splice junctions within the same RNA target.

In some embodiments designed to exclude detection of RNAs related to the target RNA, sequences are examined to identify bases that are unique to the target RNA when
20 compared to the other similar sequences from which the target is distinguished. Generally, the unique base is positioned to hybridize to the 5' end of the target-specific region of the probe oligonucleotide. In some embodiments, two adjacent bases are unique to the target compared to the related RNA. If two adjacent unique bases are available in an appropriately accessible portion of the target RNA, it is preferred that
25 these bases be used as the site around which the probe and INVADER oligonucleotides sets are designed. In some embodiments, the two unique bases are positioned such that the site of cleavage of the probe is between the two base-pairs they form with the probe. In other embodiments, one of the unique bases is in the last position of the hybridization site of the INVADER oligonucleotide (*i.e.*, it is positioned to base-pair to the penultimate
30 residue on the 3' end of the INVADER oligonucleotide).

In some embodiments, the assay is designed to include detection of RNAs that are similar, but not identical, to the target RNA. If the assay is being designed for inclusive detection, the compared sequences are examined to identify sites having complete homology. Such designs can be created to detect homologous sequences in the same species or between species. Generally, the most homologous regions are selected as hybridization targets for probe oligonucleotides. Generally, some variation can be tolerated, for example, if it is not at the base that would hybridize to the 5' end of the target-specific region of a probe. In some embodiments, variation is accommodated by the use of degenerate bases in the INVADER assay oligonucleotides (*e.g.*, mixtures of bases are used at positions within the synthesized probe, INVADER and/or stacker oligonucleotides, said mixtures selected to complement the mixture of specific bases present in the collection of related target RNAs).

iii. Oligonucleotide design

a. Target-specific regions: length and melting temperature

As described above in Section I (a) concerning the oligonucleotide design, in some embodiments, the length of the analyte-specific regions are defined by the temperature selected for running the reaction. Starting from the desired position (*e.g.*, a variant position or splice junction in a target RNA, or a site corresponding to a low free energy value in an OligoWalk analysis) an iterative procedure is used by which the length of the ASR is increased by one base pair until a calculated optimal reaction temperature (T_m plus salt correction to compensate for enzyme and any other reaction conditions effects) matching the desired reaction temperature is reached. In general probes are selected to have an ASR with a calculated T_m of about 60 °C if a stacking oligonucleotide is not used, and a T_m of about 50 to 55 °C if a stacking oligonucleotide is used (a stacking oligonucleotide typically raises the T_m of a flanking probe oligonucleotide by about 5 to 15 °C). If the position of variation or a splice junction is a starting position, then the additions are made to the 3' end of the probe. Alternatively, if the 3' end of the probe is to be positioned at the most accessible site, the additions are in the 5' direction. In some embodiments, wherein a stacker oligonucleotide is to be used, it is preferred that the

probe be designed to have a 3' base that has stable stacking interaction interface with the 5' base of the stacker oligonucleotide. The stability of coaxial stacking is highly dependent on the identity of the stacking bases. Overall, the stability trend of coaxial stacking in decreasing order is purine:purine > purine:pyrimidine \approx pyrimidine:purine > pyrimidine:pyrimidine. In other embodiments employing a stacker, a less stable stacking interaction is preferred; in such cases the probe 3' base and/or the stacker 5' base are selected to provide a less stable stacking interaction. In some embodiments, the probe 3' base and/or the stacker 5' base are selected to have a mismatch with respect to the target strand, to reduce the strength of the stacking interaction.

The same principles are also followed for INVADER oligonucleotide design. Briefly, starting from the position N, additional residues complementary to the target RNA starting from residue N-1 are then added in the upstream direction until the stability of the INVADER-target hybrid exceeds that of the probe (and therefore the planned assay reaction temperature). In preferred embodiments, the stability of the INVADER-target hybrid exceeds that of the probe by 12-15 °C. In general, INVADER oligonucleotides are selected to have a T_m near 75 °C. Software applications, such as INVADERCREATOR (Third Wave Technologies, Madison, WI) or Oligonucleotide 5.0 may be used to assist in such calculations.

If a stacking oligonucleotide is to be used, similar design principles are applied. The stacking oligonucleotide is generally designed to hybridize at the site adjacent to the 3' end of the probe oligonucleotide, such that the stacker/target helix formed can coaxially stack with the probe/target helix. The sequence is selected to have a calculated T_m of about 60 to 65 °C, with the calculation based on the use of natural bases. However, stacking oligonucleotides are generally synthesized using only 2'-O-methyl nucleotides, and consequently, have actual T_m s that are higher than calculated by about 0.8°C per base, for actual T_m s close to 75 °C.

In some embodiments, ARRESTOR oligonucleotides are included in a secondary reaction. ARRESTOR oligonucleotides are provided in a secondary reaction to sequester any remaining uncleaved probe from the primary reaction, to preclude interactions between the primary probe and the secondary target strand. ARRESTOR oligonucleotides are generally 2'-O-methylated, and comprise a portion that is

complementary to essentially all of their respective probe's target-specific region, and a portion that is complementary to at least a portion of the probe's flap regions (*e.g.*, six nucleotides, counted from the +1 base towards the 5' end of the arm).

b. Non-complementary regions

Probe 5' Arm selection

The non-complementary arm of the probe, if present, is preferably selected (by an iterative process as described above) to allow the secondary reaction to be performed at a particular reaction temperature. In the secondary reaction, the secondary probe is generally cycling, and the cleaved 5' arm (serving as an INVADER oligonucleotide) should stably bind to the secondary target strand.

INVADER oligonucleotide 3' terminal mismatch selection

In preferred embodiments, the 3' base of the INVADER oligonucleotide is not complementary to the target strand, and is selected in the following order of preference (listed as INVADER oligonucleotide 3' base/target base):

| | |
|--------------|-----------------------|
| C in target: | C/C > A/C > T/C > G/C |
| A in target: | A/A > C/A > G/A > T/A |
| G in target: | A/G > G/G > T/G > C/G |
| U in target: | C/U > A/U > T/U > G/U |

c. Folding and dimer analysis

In some embodiments, the oligonucleotides proposed for use in the INVADER assay are examined for possible inter- and intra-molecular structure formation in the absence of the target RNA. In general, it is desirable for assay probes to have fewer predicted inter- or intra molecular interactions. In some embodiments, the program OLIGO (*e.g.*, OLIGO 5.0, Molecular Biology Insights, Inc., Cascade, CO) is used for such analysis. In other embodiments, the program *mfold* is used for the analysis. In yet other embodiments, the RNAStructure program can be used for dimer analysis. The

following sections provide stepwise instructions for the use of these programs for analysis of INVADER assay oligonucleotides.

OLIGO 5.0 analysis for Probe structure and interaction prediction.

5 Analysis of INVADER oligonucleotides using OLIGO 5.0 comprises the following steps. All menu choices are shown in UPPER CASE type.

1. Launch OLIGO 5.0 and open a sequence file for each mRNA to be analyzed. This is done by using a menu to select the following

- 10 • Choose FILE - >NEW
- Paste in longest available sequence
- Choose ACCEPT & QUIT (F6)

2. Set Program settings to default

Choose FILE - >RESET - >ORIGINAL DEFAULTS

15 **3. Identify Probe Oligonucleotide**

- Select OLIGO LENGTH to be around 16 nucleotides (open the menu for this option by using ctrl-L keystrokes).
- Move the cursor indicating the 5' end of the Current Oligo until the 3' end is located at the candidate cleavage site residue.
- 20 • Choose ANALYSE - >DUPLEX FORMATION - >CURRENT OLIGO (ctrl-D) for a rough determination of the extent of dimer and hairpin formation.
- Confirm length of analyte region corresponds with desired reaction temperature [e.g., through the use of T_m calculation as described in the Optimization of Reaction Conditions, I (c) of the Detailed Description of the Invention]
- 25 • Select the "LOWER" button in OLIGO 5.0 to copy the anti-sense sequence (this will be the analyte-specific region of the actual probe oligonucleotide and is anti-sense to the RNA strand.)
- Import into a database file.

- Save to computer memory.

4. Identify INVADER Oligonucleotide

- Choose sequence adjacent to the probe oligonucleotide identified from step 3.
- Select OLIGO LENGTH to ~24 nucleotides
- 5 • Confirm length of analyte region corresponds with desired reaction temperature [e.g., through the use of T_m calculation as described in the Optimization of Reaction Conditions, I (c) of the Detailed Description of the Invention, about 75 °C for INVADER oligonucleotides). Select the "LOWER" button in OLIGO 5.0 to copy the corresponding anti-sense sequence (this will be the analyte-specific region of the actual INVADER oligonucleotide.)
- 10 • Import into a database file.
- Save to computer memory.

5. Addition of Cleaved Arm Sequence and INVADER Oligonucleotide Mismatch Sequence.

- 15 • Export the Probe oligonucleotide as Upper Primer.
- Export the INVADER oligonucleotide as Lower Primer.
- EDIT UPPER PRIMER to add in a candidate arm sequence (selected, for example, as described above).
- Check that the arm sequence does not create new secondary structures (analysis performed as described above).
- 20 • EDIT LOWER PRIMER to add in the 3' mismatched nucleotide that will overlap into the cleavage site (selected according to the guidelines for this mismatched bases, provided above).
- Select all Upper and Lower Primer boxes in the "Print/Save Options"
- 25 • PRINT ANALYSIS of Upper (Probe) and Lower (INVADER) oligonucleotides and check for lack of stable secondary structures.
- Save both mRNA sequence and oligonucleotide sequence database files before quitting the program.

Generally, oligonucleotides having detected intra-molecular formations with stabilities of less than $-6 \Delta G$ are preferred. Less stable structures represent poor substrates for CLEAVASE enzymes, and thus cleavage of such structures is less likely to contribute to background signal. Probe and INVADER oligonucleotides having less affinity for each other are more available to bind to the target, ensuring the best cycling rates.

The T_m of dimerized probes (*i.e.*, probes wherein one probe molecule is hybridized to another probe molecule) should ideally be lower than the T_m for the probe hybridized to the target, to ensure that the probes preferentially hybridize to the target sequence at the elevated temperatures at which INVADER assay reactions are generally conducted. Similarly, the T_m for the INVADER oligonucleotide hybridized either to itself or to a probe molecule should be lower than the INVADER oligonucleotide/target T_m . It is preferred that dimer T_m s (*i.e.*, Probe/Probe and Probe/INVADER oligonucleotide) be 25°C or less to ensure that they will be unlikely to form at the planned reaction temperature.

The melting temperatures for each of these complexes can be determined as described above in Optimization of Reaction Conditions, I (c) of the Detailed Description of the Invention, or by using the OLIGO software. Once RNAs sites and several candidate INVADER assay oligonucleotide sets are selected according to the process outlined above, the candidate oligonucleotide sets can be ranked according to the degree to which they comply with preferred selection rules, *e.g.*, their location on the SS-Count average plot (peak, valley, neither), and the energetic predictions of probe and INVADER oligonucleotide interactions. In some embodiments, the ranked probe sets are tested in order of rank to identify one or more sets having suitable performance in an RNA INVADER assay. In other embodiments, several of the top ranked sets (*e.g.*, two, three or more) are selected for testing, to rapidly identify one or more sets having suitable or desirable performance.

***Mfold* analysis for probe structure and interaction prediction**

Analysis of probe and INVADER oligonucleotide interactions may be performed using *mfold* for DNA provided by Michael Zuker, available through Rensselaer

Polytechnic Institute at bioinfo.math.rpi.edu/~mfold/dna/form1.cgi. The analysis is performed without changing the default ionic conditions, and with a selected temperature of 37 °C and with % suboptimality set to 75. Each sequence (*e.g.*, probe, INVADER oligonucleotide, stacker, etc.) is folded using the program to check for any unimolecular structure formation (*e.g.*, hairpins). The energies provided by *mfold* gives for unimolecular structures can be used as provided, without further calculations.

Bimolecular structure formation for a given oligonucleotide is assessed by typing in the oligonucleotide sequence (5' to 3') followed by the sequence of a small, stable hairpin forming sequence (*e.g.*, CCCCTTTTGGGGG [SEQ ID NO:707]), followed by the same oligonucleotide sequence, again listed 5' to 3'. Constraints are entered to require that these Ts remain single-stranded and the strings of Cs and Gs in this spacer are basepaired. The command "F" is used to force basepairing, while the command "P" is used to prohibit basepairing, and the positions of the forced or prohibited basepairs are counted from the 5' end. For example, if the sequence of interest is a 20-mer, then the following is entered:

F 21 0 5 [this forces the C's, C21 to C25, to base pair]

P 26 0 4 [this forces the T's, T26 to T29, to be single stranded]

F 30 0 5 [this forces the G's, G30 to G34, to base pair]

On examination of the resulting structures, the stability of each can be estimated by subtracting the stability (*i.e.*, the thermodynamic measures) of the central spacer hairpin from the total result (*i.e.*, Thermodynamics of possible structure = *mfold* structure thermodynamics – core hairpin thermodynamics). For convenience, in some embodiments, any nearest neighbor interactions between the central hairpin and dimers formed by the test sequence are ignored for this calculation; a more accurate analysis would require consideration of this interaction. The core hairpin formed by CCCCTTTTGGGGG (SEQ ID NO:707) has the following thermodynamics: $\Delta G = -5.3$; $\Delta H = -37.8$; $\Delta S = -104.8$.

The process can be demonstrated using the following probe sequence: 5'-CCCTATCTTTAAAGTTTTTAAAAAGTTTGA-3' (SEQ ID NO:708). The oligonucleotide sequence is examined by *mfold* analysis for bimolecular structures using the following steps.

1- In *mfold* sequence box type:

```
CCCTATCTTTAAAGTTTTTAAAAAGTTTGACCCCCTTTTGGGGGGCCC
TATCTTTAAAGTTTTTAAAAAGTTTGA (SEQ ID NO:137)
```

5 2- In the constraint box type:

```
P 36 0 4
F 31 0 5
F 40 0 5
```

10 Results (showing one):

```
Structure 1
dG =  -14.2    dH =  -150.5    dS =  -439.5    Tm =  69.3
CCCTATCTTT |G          G          -----    T
          AAA TTTTAAAAA TTTGA          CCCCC T
15          TTT AAAAATTTT AAATT          GGGGG T
-----AG ^G          G          TCTATCCC    T
```

To evaluate the stability of the duplex:

```
20 CCCTATCTTT |G          G
          AAA TTTTAAAAA TTTGA
          TTT AAAAATTTT AAATT
-----AG ^G          G          TCTATCCC
```

the thermodynamic values for the hairpin alone are subtracted from the values for the
25 complete structure:

$$\Delta G = -14.2 - (-5.3) = -8.9,$$

$$\Delta H = -150.5 - (-37.8) = -112.7,$$

$$\Delta S = -439.5 - (-104.8) = -334.7,$$

Using a calculation wherein $T_m (^{\circ}\text{C}) = \{\Delta H / [\Delta S + R \ln (CT/4)]\} - 273.15$, wherein R is
30 the gas constant 1.987 (cal/K.mol), ln is the natural log, and CT is the total single strand
concentration in Molar, this results in a calculated T_m of 46.1 $^{\circ}\text{C}$ for the non-hairpin
portion of the structure.

The above method is not limited to the use of the core hairpin sequence
CCCCCTTTTGGGGG but rather any stable hairpin sequences can be used. For
35 example, CGCGCGGAACGCGCG (SEQ ID NO:138) or CCCGGGTTTTCCCGGG

(SEQ ID NO:139). However, if a different hairpin sequence is used, one needs to calculate its stability using *mfold* and use its thermodynamics in the subsequent calculations.

5 RNAStructure for oligonucleotide interaction prediction

Dimer formation can also be evaluated using the RNAStructure program. Unlike *mfold*, RNAStructure allows the calculation of all possible oligonucleotide-oligonucleotide interactions and provides an output .ct file. One can then view the structures using any .ct viewing program such as RNAStructure or RNAvis (1997, P.

10 Rijk, University of Antwerp (UIA), available on the Internet at rna.uia.ac.be/rnavis) and evaluate the stability of any dimer formation using the nearest-neighbor model (Borer et al., 1974) and DNA nearest-neighbor parameters (Allawi & SantaLucia, 1997).

For example, to evaluate the propensity of the sequence 5'

AGGCGCACCAATTTGGTGTT 3' (SEQ ID NO:140) for dimer formation using the

15 DNA Fold Intermolecular module of RNAStructure, the sequence is saved into a file (e.g., probe.seq) and the following parameters are set:

Sequence file 1: probe.seq

Sequence file 2: probe.seq

CT file: dimer.ct

20 Max % Energy difference: 50

Max number of structures: 20

Window size: do not change

After the calculation is done, one can view the resulting .ct file using the "view" module
25 of RNAStructure. Generally, there will be several structures within the .ct file. The view module is used to view them individually. One of the dimers that the test sequence, above, can form according to RNAStructure is:

AGGCG TT
 CACCAATTTGGTG
 GTGGTTTAACCAC
 TT GCGGA

5

According to the nearest-neighbor model (*i.e.*, using DNA nearest-neighbor and mismatch parameters [Allawi & SantaLucia, 1997]), the stability of this duplex in 1M NaCl and at a probe concentration of 100 μ M is:

$$\Delta G^{\circ}_{37} = -10.07$$

10 $\Delta H = -87.6$

$$\Delta S = -250.1$$

$$T_m = 50.1^{\circ}\text{C}$$

By changing the identities of Sequence Files 1 & 2, RNAstructure can be used to
 15 evaluate the possibility of any dimer formation between pairs of all of the DNA oligonucleotides present in an INVADER assay reaction.

iv. Assay performance evaluation

Probe sets selected according to the guidelines provided above can be tested in the
 20 INVADER assay to evaluate performance. While the oligonucleotides are designed to perform at or near a particular desired reaction temperature, the best performance for a given design may not be precisely at the intended temperature. Thus, in evaluating any new INVADER assay probe set, it can be helpful to examine the performance in the INVADER assay conducted at several different reaction temperatures, over a range of
 25 about 10 to 15 $^{\circ}\text{C}$, centered around the designed temperature. For convenience, temperature optimization can be performed on a temperature gradient thermocycler with a fixed amount of RNA (*e.g.*, 2.5 amoles of an *in vitro* transcript per reaction); and for a fixed amount of time (*e.g.*, 1 hour each for Primary and Secondary reactions). The temperature gradient test will reveal the temperature at which the designed probe set
 30 produces the best performance (*e.g.*, the highest level of target-specific signal compared

to background signal, generally expressed as a multiple of the zero-target background signal, or "fold over zero").

The results can be examined to see how close the measured temperature optimum is to the intended temperature of operation. In some embodiments, it is desirable to have probe sets that operate at or near a pre-selected temperature. If the measured temperature optimum is higher than the desired reaction temperature, a probe design can be altered in ways that tend to reduce the probe/target T_m (e.g., shortened by one or more bases, or altered to contain one or more mismatched bases). In some embodiments, wherein a stacker oligonucleotide is not used, wherein the reaction temperature is more than 7 °C above the desired reaction temperature, and wherein the performance (e.g., the fold over zero) is acceptable, use of a 3' mismatch on the probe oligonucleotide is likely to lower the reaction temperature without otherwise altering the assay performance.

An LOD determination can be made by performing reactions on varying amounts of target RNA (e.g., an *in vitro* transcript control RNA of known concentration). In preferred embodiments, a designed assay has an LOD of less than 0.05 attomole. In particularly preferred embodiments, a designed assay has an LOD of less than 0.01 attomole. It is contemplated that the same guideline provided above for reducing the LOD of a designed assay may be used for the purpose of raising the LOD of a designed assay, i.e., to make it LESS sensitive to the presence of a target RNA. For example, it may be desirable to detect an abundant RNA and a rare RNA in the same reaction. In such a reaction, it may be desirable to attenuate the signal generated for the abundant RNA so that it does not overwhelm the signal from the rarer species. In some embodiments this may be done by designing probe sets for reduced signal generation, e.g., an LOD of at least (not less than) 0.5 attomoles. In some embodiments, a single step INVADER assay may be used for detection of abundant targets in a sample, while sequential INVADER reactions to amplify signal, as described in Section II, may be used for less abundant analytes in the same sample. In preferred embodiments, the single step and the sequential INVADER assay reactions for the different analytes are performed in a single reaction.

In some embodiments, time course reactions are run, wherein the accumulation of signal for a known amount of target is measured for reactions run for different lengths of

time. This measurement will establish the linear ranges, *i.e.*, the ranges in which accurate quantitative measurements can be made using a given assay design, with respect to time and starting target RNA level.

5 **v. Design and assay optimization**

Some designed assays may not meet the preferred performance criteria described above. A number of variations on the performance of INVADER assay reactions have been described herein. In optimizing performance of the INVADER assay for the detection of RNA targets, these variations may be used alone or in combination. For
10 example, in some embodiments, a stacker oligonucleotide is employed. While not limiting the present invention to any particular mechanism of action, in some embodiments, a stacker oligonucleotide may enhance performance of an assay by altering the hybridization characteristics (*e.g.*, T_m) of a probe or an INVADER oligonucleotide. In some embodiments, a stacker oligonucleotide may increase performance by enabling
15 the use of a shorter probe. In other embodiments, a stacker oligonucleotide may enhance performance by altering the folded structure of the target nucleic acid. In yet other embodiments, the enhancing activity of the stacker oligonucleotide may involve these and other mechanisms in combination.

In other embodiments, the target site may be shifted. In some embodiments,
20 reactions are optimized by testing multiple probe sets that shift along a suspected accessible site. In preferred embodiments, such probe sets shift along the accessible site in one to two base increments. In embodiments wherein accessible sites have previously been predicted only by computer analysis, physical detection of the accessible sites may be employed to optimize a probe set design. In preferred embodiments, the RT-ROL
25 method of detecting accessible sites is employed. In some embodiments, optimization of a probe set design may require shifting of the target site to a newly identified accessible site.

In some embodiments, *e.g.*, wherein an accessible site has been identified yet probe set performance is low, a change in the design of a probe 5' arm may improve
30 assay performance without altering the site targeted. In other embodiments, altering the length of an ARRESTOR oligonucleotide (*e.g.*, increasing the length of the portion that is

complementary to the 5' arm region of the probe) may reduce background signal, thus increasing the probe set performance.

Other variations on oligonucleotide design may be employed to alter performance in an assay. Some modifications may be employed to shift the ideal operating

5 temperature of a probe set design into a preferred temperature range. For example, the use of shorter oligonucleotides and the incorporation of mismatches generally act to reduce the T_m s, and thus reduce the ideal operating temperatures, of designed oligonucleotides. Conversely, the use of longer oligonucleotides and the employment of stacking oligonucleotides generally act to increase the T_m s, and thus increase the ideal
10 operating temperatures of the designed oligonucleotides.

Other modifications may be employed to alter other aspects of oligonucleotide performance in an assay. For example, the use of base analogs or modified bases can alter enzyme recognition of the oligonucleotide. In some embodiments, such modified bases are used to protect a region of an oligonucleotide from nuclease cleavage. In other
15 embodiments, modified bases are used to affect the ability of an oligonucleotide to participate as a member of a cleavage structure that is not in a position to be cleaved (*e.g.*, to serve as an INVADER oligonucleotide to enable cleavage of a probe). These modified bases may be referred to as "blocker" or "blocking" modifications. In some embodiments, assay oligonucleotides incorporate 2'-O-methyl modifications. In other
20 embodiments, assay oligonucleotides incorporate 3' terminal modifications (*e.g.*, NH_2 , 3' hexanol, 3' phosphate, 3' biotin).

In yet other embodiments, the components of the reaction may be altered to affect assay performance. For example, oligonucleotide concentrations may be varied.

Oligonucleotide concentrations can affect multiple aspects of the reaction. Since melting
25 temperatures of complexes are partly a function of the concentrations of the components of the complex, variation of the concentrations of the oligonucleotide components can be used as one facet of reaction optimization. In the methods of the present invention, ARRESTOR oligonucleotides may be used to modulate the availability of the primary probe oligonucleotides in an INVADER assay reaction. In some embodiments, an
30 ARRESTOR oligonucleotide may be excluded. Other reaction components may also be varied, including enzyme concentration, salt and divalent ion concentration and identity.

VI. Kits for performing the RNA INVADER Assay

In some embodiments, the present invention provides kits comprising one or more
5 of the components necessary for practicing the present invention. For example, the
present invention provides kits for storing or delivering the enzymes of the present
invention and/or the reaction components necessary to practice a cleavage assay (e.g., the
INVADER assay). The kit may include any and all components necessary or desired for
the enzymes or assays including, but not limited to, the reagents themselves, buffers,
10 control reagents (e.g., tissue samples, positive and negative control target
oligonucleotides, etc.), solid supports, labels, written and/or pictorial instructions and
product information, inhibitors, labeling and/or detection reagents, package
environmental controls (e.g., ice, desiccants, etc.), and the like. In some embodiments,
the kits provide a sub-set of the required components, wherein it is expected that the user
15 will supply the remaining components. In some embodiments, the kits comprise two or
more separate containers wherein each container houses a subset of the components to be
delivered. For example, a first container (e.g., box) may contain an enzyme (e.g.,
structure specific cleavage enzyme in a suitable storage buffer and container), while a
second box may contain oligonucleotides (e.g., INVADER oligonucleotides, probe
20 oligonucleotides, control target oligonucleotides, etc.). In some embodiments one or
more the reaction components may be provided in a predispensed format (*i.e.*,
premeasured for use in a step of the procedure without re-measurement or re-dispensing).
In some embodiments, selected reaction components are mixed and predispensed
together. In preferred embodiments, predispensed reaction components are predispensed
25 and are provided in a reaction vessel (including but not limited to a reaction tube or a
well, as in, *e.g.*, a microtiter plate). In particularly preferred embodiments, predispensed
reaction components are dried down (*e.g.*, desiccated or lyophilized) in a reaction vessel.

Additionally, in some embodiments, the present invention provides methods of
delivering kits or reagents to customers for use in the methods of the present invention.
30 The methods of the present invention are not limited to a particular group of customers.
Indeed, the methods of the present invention find use in the providing of kits or reagents

to customers in many sectors of the biological and medical community, including, but not limited to customers in academic research labs, customers in the biotechnology and medical industries, and customers in governmental labs. The methods of the present invention provide for all aspects of providing the kits or reagents to the customers, including, but not limited to, marketing, sales, delivery, and technical support.

In some embodiments of the present invention, quality control (QC) and/or quality assurance (QA) experiments are conducted prior to delivery of the kits or reagents to customers. Such QC and QA techniques typically involve testing the reagents in experiments similar to the intended commercial uses (*e.g.*, using assays similar to those described herein). Testing may include experiments to determine shelf life of products and their ability to withstand a wide range of solution and/or reaction conditions (*e.g.*, temperature, pH, light, etc.).

In some embodiments of the present invention, the compositions and/or methods of the present invention are disclosed and/or demonstrated to customers prior to sale (*e.g.*, through printed or web-based advertising, demonstrations, etc.) indicating the use or functionality of the present invention or components of the present invention. However, in some embodiments, customers are not informed of the presence or use of one or more components in the product being sold. In such embodiments, sales are developed, for example, through the improved and/or desired function of the product (*e.g.*, kit) rather than through knowledge of why or how it works (*i.e.*, the user need not know the components of kits or reaction mixtures). Thus, the present invention contemplates making kits, reagents, or assays available to users, whether or not the user has knowledge of the components or workings of the system.

Accordingly, in some embodiments, sales and marketing efforts present information about the novel and/or improved properties of the methods and compositions of the present invention. In other embodiments, such mechanistic information is withheld from marketing materials. In some embodiments, customers are surveyed to obtain information about the type of assay components or delivery systems that most suits their needs. Such information is useful in the design of the components of the kit and the design of marketing efforts.

VII. The INVADER Assay for Direct Detection and Measurement of Specific RNA Analytes.

The following section provides a few illustrative examples of mRNAs that may be
5 detected or measured using the methods, compositions and systems of the present
invention.

HOUSEKEEPING CONTROLS

10 RNAs that are generally present in predicable or invariant amounts in test samples
provide useful control targets for detection assays. These controls can be useful in
several ways, including but not limited to providing confirmation of the proper function
of an assay, and as a standard against which a test result for another RNA can be
compared or measured to aid in interpretation of a result. mRNAs for the following
genes find particular use in the methods of the present invention.

15 Human Ubiquitin and Mouse/Rat Ubiquitin

The ubiquitin system is a major pathway for selective protein degradation.
Degradation by this system is instrumental in a variety of cellular functions such as DNA
repair, cell cycle progression, signal transduction, transcription, and antigen presentation.
The ubiquitin pathway also eliminates proteins that are misfolded, misplaced, or that are
20 in other ways abnormal. This pathway requires the covalent attachment of ubiquitin
(E1), a highly conserved 76 amino acid protein, to defined lysine residues of substrate
proteins.

Human, rat and mouse glyceraldehyde-3-phosphate dehydrogenase (GAPDH)

25 GAPDH is an important enzyme in the glycolysis and gluconeogenesis pathways.
This homotetrameric enzyme catalyzes the oxidative phosphorylation of D-
glyceraldehyde-3-phosphate to 1,3-diphosphoglycerate in the presence of cofactor and
inorganic phosphate. A variety of diverse biological properties of GAPDH have been
reported. These include functions in endocytosis, mRNA regulation, tRNA export, DNA
30 replication, DNA repair, and neuronal apoptosis.

CYTOKINES

A growing family of regulatory proteins that deliver signals between cells of the immune system has been identified. Called cytokines, these proteins have been found to control the growth and development, and bioactivities, of cells of the hematopoietic and immune systems. Cytokines exhibit a wide range of biological activities with target cells from bone marrow, peripheral blood, fetal liver, and other lymphoid or hematopoietic organs. The present invention describes methods for the detection of expression of cytokines, including but not limited to of the exemplary members of the cytokine family listed below.

Human Oncostatin M

Oncostatin M is a secreted single-chain polypeptide cytokine that regulates the growth of certain tumor-derived and normal cell lines. A number of cell types have been found to bind the oncostatin M protein. While it has been shown to inhibit proliferation of a number of tumor cell types, it has also been implicated in stimulating proliferation of Kaposi's sarcoma cells.

Human transforming growth factor-beta (TGF- β)

Transforming growth factor-beta (TGF-beta) is a member of a family of structurally-related cytokines that elicit a variety of responses, including growth, differentiation, and morphogenesis, in many different cell types. In vertebrates, at least five different forms of TGF-beta, termed TGF-beta1 to TGF-beta5, have been identified; they all share a high degree (60%-80%) of amino-acid sequence identity. While TGF-beta1 was initially characterized by its ability to induce anchorage-independent growth of normal rat kidney cells, its effects on most cell types are anti-mitogenic. It is strongly growth-inhibitory for many types of cells, including both normal and transformed epithelial, endothelial, fibroblast, neuronal, lymphoid, and hematopoietic cells. In addition, TGF-beta plays a central role in regulating the formation of extracellular matrix and cell-matrix adhesion processes.

Human monocyte chemoattractant protein-1 (MCP-1)

Within this family of cytokines, an emerging group of chemotactic cytokines, also called chemokines or intercrines, has been identified. Two subfamilies of chemokines have been recognized, alpha and beta, based on chromosomal location and the arrangement of the cysteine residues.

The human genes encoding the beta subfamily proteins are located on chromosome 17 (their mouse counterparts are clustered on mouse chromosome 11, which is the counterpart of human chromosome 17). Homology in the beta subfamily ranges from 28-45% intraspecies, from 25-55% interspecies. An exemplary member is the human protein MCP-1 (monocyte chemoattractant protein-1). MCP-1 exerts several effects specifically on monocytes. It is a potent chemoattractant for human monocytes *in vitro* and can stimulate an increase in cytosolic free calcium and the respiratory burst in monocytes. MCP-1 has been reported to activate monocyte-mediated tumoricidal activity, as well as to induce tumoricidal activity. MCP-1 has been implicated as an important factor in mediating monocytic infiltration of tissues inflammatory processes such as rheumatoid arthritis and alveolitis. The factor may also play a fundamental role in the recruitment of monocyte-macrophages into developing atherosclerotic lesions.

Human tumor necrosis factor alpha (TNF- α)

Tumor necrosis factor alpha (TNF-alpha also cachectin) is an important cytokine that plays a role in host defense. The cytokine is produced primarily in macrophages and monocytes in response to infection, invasion, injury, or inflammation. Some examples of inducers of TNF-alpha include bacterial endotoxins, bacteria, viruses, lipopolysaccharide (LPS) and cytokines including GM-CSF, IL-1, IL-2 and IFN-gamma.

Despite the protective effects of the cytokine, overexpression of TNF-alpha often results in disease states, particularly in infectious, inflammatory and autoimmune diseases. This process may involve the apoptotic pathways. High levels of plasma TNF-alpha have been found in infectious diseases such as sepsis syndrome, bacterial meningitis, cerebral malaria, and AIDS; autoimmune diseases such as rheumatoid arthritis, inflammatory bowel disease (including Crohn's disease), sarcoidosis, multiple sclerosis, Kawasaki syndrome, graft-versus-host disease and transplant (allograft)

rejection; and organ failure conditions such as adult respiratory distress syndrome, congestive heart failure, acute liver failure and myocardial infarction. Other diseases in which TNF-alpha is involved include asthma, brain injury following ischemia, non-insulin-dependent diabetes mellitus, insulin-dependent diabetes mellitus, hepatitis, atopic dermatitis, and pancreatitis. Further, inhibitors of TNF-alpha have been suggested to be useful for cancer prevention. Elevated TNF-alpha expression may also play a role in obesity. TNF-alpha was found to be expressed in human adipocytes and increased expression, in general, correlated with obesity.

10 **Human interleukin-6 (IL-6)**

IL-6 is the standardized name of a cytokine called B lymphocyte differentiating factor, interferon beta2, 26 Kd protein, hybridoma/plasmacytoma growth factor, hepatocyte stimulating factor, etc.

IL-6 induces activated B cells to be differentiated into antibody forming cells.

15 For T cells, IL-6 induces T cells stimulated by mitogens to produce IL-2 and induces the expression of IL-2 receptor on a certain T cell line or thymocytes. For blood forming cells, IL-6 induces the growth of blood forming stem cells synergistically in the presence of IL-3. Furthermore, recently, it was reported that IL-6 acted like thrombopoietin.

IL-6 is produced by various cells. It is produced by lymphocytes and is also
20 produced by human fibroblasts stimulated by Poly (I)-Poly (C) and cycloheximide. Murine IL-6 is produced in mouse cells, which are stimulated by Poly (A)-Poly (U). Inducers for stimulation are diverse, and include known cytokines such as IL-1, TNF and IFN-beta, growth factors such as PDGF and TGF-beta, LPS, PMA, PHA, cholera toxin, etc. Moreover, it is reported that human vascular endothelial cells, macrophages, human
25 glioblastomas, etc. also produce IL-6. Furthermore, it is also known that the productivity can be further enhanced by stimulating cells using an inducer and subsequently treating the cells by a metabolic inhibitor such as verapamil, cycloheximide or actinomycin D, etc.

Human interleukin 1beta (IL-1 β)

Interleukin-1 (IL-1) is important to the activation of T and B lymphocytes and mediates many inflammatory processes. Two distinct forms of IL-1 have been isolated and expressed; termed IL-1beta and IL-1alpha. IL-1beta is the predominant form produced by human monocytes both at the mRNA and protein level. The two forms of human IL-1 share only 26% amino acid homology. Despite their distinct polypeptide sequences, the two forms of IL-1 have structural similarities, in that the amino acid homology is confined to discrete regions of the IL-1 molecule. The two forms of IL-1 also possess identical biological properties, including induction of fever, slow wave sleep, and neutrophilia, T- and B-lymphocyte activation, fibroblast proliferation, cytotoxicity for certain cells, induction of collagenases, synthesis of hepatic acute phase proteins, and increased production of colony stimulating factors and collagen. IL-1 also activates endothelial cells, resulting in increased leukocyte adhesiveness, PGI₂ and PGE₂ (prostaglandins) release, and synthesis of platelet activating factor, procoagulant activity, and a plasminogen activator inhibitor. Clearly, IL-1 plays a central role in local and systemic host responses. Because many of the biological effects of IL-1 are produced at picomolar (pg) concentrations *in vivo*, IL-1 production is likely a fundamental characteristic of host defense mechanisms.

Human interleukin 2 (IL-2)

Interleukin-2 (IL-2) is the main growth factor of T lymphocytes. By regulating T helper lymphocyte activity IL-2 increases the humoral and cellular immune responses. By stimulating cytotoxic CD8 T cells and NK cells, this cytokine participates in the defense mechanisms against tumors and viral infections. IL-2 is used in therapy against metastatic melanoma and renal adenocarcinoma. IL-2 is used in clinical trials in many forms of cancer. It is also used in HIV infected patients and leads to a significant increase in CD4 counts. Human IL-2 is a protein of 133 amino acids (aa) composed of four alpha helices connected by loops of various length, its tridimensional structure has been established. IL-2R is composed of three chains alpha, beta and gamma. IL2Ralpha controls the affinity of the receptor IL-2Rbeta and IL-2Rgamma are responsible for IL-2 signal transduction. The different molecular areas of IL-2 interacting with the three

chains of the IL-2 R have been defined. More specifically it has been determined that a helix A as well as the NH₂ terminal area of IL-2 (residues 1 to 30) control the interactions IL-2/IL-2Rbeta.

5 **Human interleukin 8 (IL-8)**

Human IL-8 is a cytokine that has variously been called neutrophil-activating protein, neutrophil chemotactic factor (NCF) and T-cell chemotactic factor. IL-8 can be secreted by several types of cells upon appropriate stimulation. IL-8 is secreted by activated monocytes and macrophages as well as by embryonic fibroblasts.

10 IL-8 is known to induce neutrophil migration and to activate functions of neutrophils such as degranulation, release of superoxide anion and adhesion to the endothelial cell monolayer. There are a number of conditions that are known to involve leukocyte infiltration into lesions. These include pulmonary diseases such as pulmonary cystic fibrosis, idiopathic pulmonary fibrosis, adult respiratory distress syndrome, 15 sarcoidosis and empyema; dermal diseases such as psoriasis, rheumatoid arthritis; and inflammatory bowel disease (Crohn's Disease).

The amino acid sequence characterizing human IL-8 was described by Matsushima, et al. in PCT application WO89/08665. More recently, it was shown that monocyte-derived IL-8 was evidently variably processed at the N-terminus and that the 20 IL-8 originally disclosed by Matsushima et al. was accompanied by two forms of the factor which had seven or five additional amino acids at the N-terminus (Yoshimura, *et al.*, Mol Immunol 26:87 [1989]). The longest form accounted for about 8%, the next longest form for about 47%, and the shortest form for about 45% of the total IL-8 derived from monocytes.

25

Human interleukin 10 (IL-10)

Interleukin-10 (IL-10), a recently discovered lymphokine, was originally described as an inhibitor of interferon-gamma synthesis and is postulated as a major mediator of the humoral class of immune response. Two classes of often mutually 30 exclusive immune responses are the humoral (antibody-mediated) and the delayed-type hypersensitivity.

It is postulated that these two differing immune responses may arise from two types of helper T-cell clones, namely Th1 and Th2 helper T-cells, which demonstrate distinct cytokine secretion patterns. Mouse Th1 cell clones secrete interferon-gamma, and IL-2 and preferentially induce the delayed-type hypersensitivity response while Th-2 cell clones secrete IL-4, IL-5 and IL-10 and provide support for the humoral responses. The contrast in immune response could result because interferon-gamma secreted by the Th1 cell clones inhibits Th2 clone proliferation in vitro, while IL-10 secreted by the Th2 cell clones inhibits cytokine secretion by the Th1 cell clones. Thus the two T-helper cell types may be mutually inhibitory and may provide the underpinning for the two dissimilar immune responses.

IL-10 has been cloned and sequenced from both murine and human T cells. Both sequences contain an open reading frame encoding a polypeptide of 178 amino acids with an N-terminal hydrophobic leader sequence of 18 amino acids, and have an amino acid sequence homology of 73%.

15

Human interleukin 4 (IL-4)

Interleukin-4 (IL-4, also known as B cell stimulating factor, or BSF-1) was originally characterized by its ability to stimulate the proliferation of B cells in response to low concentrations of antibodies directed to surface immunoglobulin. More recently, IL-4 has been shown to possess a far broader spectrum of biological activities, including growth co-stimulation of T cells, mast cells, granulocytes, megakaryocytes, and erythrocytes. In addition, IL-4 stimulates the proliferation of several IL-2- and IL-3-dependent cell lines, induces the expression of class II major histocompatibility complex molecules on resting B cells, and enhances the secretion of IgE and IgG1 isotypes by stimulated B cells. Both murine and human IL-4 have been definitively characterized by recombinant DNA technology and by purification to homogeneity of the natural murine protein.

The biological activities of IL-4 are mediated by specific cell surface receptors for IL-4 that are expressed on primary cells and in vitro cell lines of mammalian origin. IL-4 binds to the receptor, which then transduces a biological signal to various immune effector cells.

Human interferon gamma (IFN- γ)

Like the interleukins, interferons belong to the class of the cytokines and are listed in various classes: interferon-alpha, interferon-beta, interferon-gamma, interferon-omega and interferon-tau. Interferon-gamma is a glycoprotein, the amino acid sequence of which has been known since 1982. In the mature condition the interferon-gamma has 143 amino acids and a molecular weight of 63 to 73 kilodaltons.

The tertiary and quaternary structure of the non-glycosylised protein was clarified in 1991. According to this, interferon-gamma exists as a homodimer, the monomers being orientated in contrary directions in such a way that the C-terminal end of one monomer is located in the vicinity of the N-terminal end of the other monomer. Each of these monomers in all has six alpha-helices. Interferon-gamma is also called immunointerferon, as it has non-specific antiviral, antiproliferative and in particular immunomodulatory effects. Its production in T-helper-lymphocytes is stimulated by mitogens and antigens. The effect of the expressed interferon-gamma has not yet been precisely clarified, but is subject to intensive research. In particular, interferon-gamma leads to the activation of macrophages and to the synthesis of histocompatibility antigens of the class 2. *In vitro*, the activity of interferon-gamma is normally determined as a reduction in the virus-induced cytopathic effect, which arises from treatment with interferon-gamma. Due to its antigen-non-specific antiviral, antiproliferative and immunomodulatory activity it is suitable as a human therapeutic agent, for example of kidney tumours and chronic granulomatosis.

CYTOCHROME P450s

The term cytochrome P-450 refers to a family of enzymes (located on the endoplasmic reticulum, with high concentrations of the proteins in the cells of the liver and small intestine) that are primarily responsible for the metabolism of xenobiotics such as drugs, carcinogens and environmental chemicals, as well as several classes of endobiotics such as steroids and prostaglandins. Members of the cytochrome P450 family are present in varying levels and their expression and activities are controlled by variables such as chemical environment, sex, developmental stage, nutrition and age.

More than 200 cytochrome P450 genes have been identified. There are multiple forms of these P450 genes and each of the individual forms exhibit degrees of specificity towards individual chemicals in the above classes of compounds. In some cases, a substrate, whether it be drug or carcinogen, is metabolized by more than one of the cytochromes P450. Genetic polymorphisms of cytochromes P450 result in phenotypically-distinct subpopulations that differ in their ability to perform biotransformations of particular drugs and other chemical compounds.

The present invention provides methods for the detection cytochrome P450 mRNAs, including but not limited to, Human CYP 1A1, Human CYP 1A2, Human CYP 2B1, Human CYP 2B2, Human CYP 2B6, Human CYP 2C19, Human CYP 2C9, Human CYP 2D6, Human CYP 3A4, Human CYP 3A5, Human CYP 3A7, Rat CYP 2E1, Rat CYP 3A1, Rat CYP 3A2, Rat CYP 4A1, Rat CYP 4A2 and Rat CYP 4A3.

EXAMPLES

The following examples serve to illustrate certain preferred embodiments and aspects of the present invention and are not to be construed as limiting the scope thereof.

5 In the disclosure which follows, the following abbreviations apply: Afu (*Archaeoglobus fulgidus*); Mth (*Methanobacterium thermoautotrophicum*); Mja (*Methanococcus jannaschii*); Pfu (*Pyrococcus furiosus*); Pwo (*Pyrococcus woesei*); Taq (*Thermus aquaticus*); Taq DNAP, DNAP_{Taq}, and Taq Pol I (*T. aquaticus* DNA polymerase I); DNAP_{Stf} (the Stoffel fragment of DNAP_{Taq}); DNA_{PEc1} (*E. coli* DNA polymerase I); Tth (*Thermus thermophilus*); Ex. (Example); Fig. (Figure); °C (degrees Centigrade); *g* (gravitational field); hr (hour); min (minute); *oligo* (oligonucleotide); rxn (reaction); vol (volume); w/v (weight to volume); v/v (volume to volume); BSA (bovine serum albumin); CTAB (cetyltrimethylammonium bromide); HPLC (high pressure liquid chromatography); DNA (deoxyribonucleic acid); p (plasmid); μ l (microliters); ml (milliliters); μ g (micrograms); mg (milligrams); M (molar); mM (milliMolar); μ M (microMolar); pmoles (picomoles); amoles (attomoles); zmoles (zeptomoles); nm (nanometers); kdal (kilodaltons); OD (optical density); EDTA (ethylene diamine tetra-acetic acid); FITC (fluorescein isothiocyanate); SDS (sodium dodecyl sulfate); NaPO₄ (sodium phosphate); NP-40 (Nonidet P-40); Tris (tris(hydroxymethyl)-aminomethane); PMSF (phenylmethylsulfonylfluoride); TBE (Tris-Borate-EDTA, *i.e.*, Tris buffer titrated with boric acid rather than HCl and containing EDTA); PBS (phosphate buffered saline); PPBS (phosphate buffered saline containing 1 mM PMSF); PAGE (polyacrylamide gel electrophoresis); Tween (polyoxyethylene-sorbitan); ATCC (American Type Culture Collection, Rockville, MD); Coriell (Coriell Cell Repositories, Camden, NJ); DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen, Braunschweig, Germany); Ambion (Ambion, Inc., Austin, TX); Boehringer (Boehringer Mannheim Biochemical, Indianapolis, IN); MJ Research (MJ Research, Watertown, MA); Sigma (Sigma Chemical Company, St. Louis, MO); Dynal (Dynal A.S., Oslo, Norway); Gull (Gull Laboratories, Salt Lake City, UT); Epicentre (Epicentre Technologies, Madison, WI); Lampire (Biological Labs., Inc., Coopersberg, PA); MJ Research (MJ Research, Watertown, MA); National Biosciences (National Biosciences,

10
15
20
25
30

Plymouth, MN); NEB (New England Biolabs, Beverly, MA); Novagen (Novagen, Inc., Madison, WI); ; Promega (Promega, Corp., Madison, WI); Stratagene (Stratagene Cloning Systems, La Jolla, CA); Clontech (Clontech, Palo Alto, CA) Pharmacia (Pharmacia, Piscataway, NJ); Milton Roy (Milton Roy, Rochester, NY); Amersham (Amersham International, Chicago, IL); and USB (U.S. Biochemical, Cleveland, OH).
5 Glen Research (Glen Research, Sterling, VA); Coriell (Coriell Cell Repositories, Camden, NJ); Gentra (Gentra, Minneapolis, MN); Third Wave Technologies (Third Wave Technologies, Madison, WI); PerSeptive Biosystems (PerSeptive Biosystems, Framington, MA); Microsoft (Microsoft, Redmond, WA); Qiagen (Qiagen, Valencia, CA); Molecular Probes (Molecular Probes, Eugene, OR); VWR (VWR Scientific,);
10 Advanced Biotechnologies (Advanced Biotechnologies, INC., Columbia, MD); and Perkin Elmer (also known as PE Biosystems and Applied Biosystems, Foster City, CA).

EXAMPLE 1

Rapid screening of colonies for 5' nuclease activity

The native 5' nucleases and the enzymes of the present invention can be tested directly for a variety of functions. These include, but are not limited to, 5' nuclease activity on RNA or DNA targets and background specificity using alternative substrates
20 representing structures that may be present in a target detection reaction. Examples of nucleic acid molecules having suitable test structures are shown schematically in Figures 18A-D and Figures 21-24. The screening techniques described below were developed to quickly and efficiently characterize 5' nucleases and to determine whether the new 5' nucleases have any improved or desired activities. Enzymes that show improved cycling
25 rates on RNA or DNA targets, or that result in reduced target-independent cleavage merit more thorough investigation. In general, the modified proteins developed by random mutagenesis were tested by rapid colony screen on the substrates shown in Figures 18A and 18B. A rapid protein extraction was then done, and a test of activity on alternative structures, (*e.g.*, as shown in Figures 18C-D) was performed using the protein extract.
30 Either the initial screen, or further screening and characterization of enzymes for improved activity may be performed using other cleavage complexes, such as those

diagrammed in Figures 21-24. It is not intended that the scope of the invention be limited by the particular sequences used to form such test cleavage structures. One skilled in the art would understand how to design and create comparable nucleic acids to form analogous structures for rapid screening.

5 This order of testing may be chosen to reduce the number of tests overall, to save time and reagents. The order of testing for enzyme function is not intended to be a limitation on the present invention. Those mutants that showed reasonable cycling rates with the RNA or DNA targets may then be cultured overnight, and a rapid protein extraction done. Alternatively, any subset or all of the cleavage tests may be done at the
10 same time.

For convenience, each type of rapid screen may be done on a separate microtiter plate. For example, one plate may be set up to test RNA INVADER activity, one plate set up to test for DNA INVADER activity. As many as 90 different colonies may be screened on one plate. The colonies screened can be from a variety of sources, such as
15 clones of unaltered (native) 5' nucleases, from one mutagenesis reaction (*e.g.*, many colonies from a single plate) or from a variety of reactions (colonies selected from multiple plates).

Ideally, positive and negative controls should be run on the same plate as the mutants, using the same preparation of reagents. One example of a good positive control
20 is a colony containing the unmodified enzyme, or a previously modified enzyme whose activity is to be compared to new mutants. For example, if a mutagenesis reaction is performed on the Taq DN RX HT construct (described below), the unmodified Taq DN RX HT construct would be chosen as the standard for comparing the effects of mutagenesis on enzymatic activity. Additional control enzymes may also be incorporated
25 into the rapid screening test. For example, Tth DN RX HT (described below; unless otherwise specified, the TaqPol and TthPol enzymes of the following discussion refer to the DN RX HT derivative) may also be included as a standard for enzymatic activity along with the Taq DN RX HT. This would allow a comparison of any altered enzymes to two known enzymes having different activities. A negative control should also be run
30 to determine the background reaction levels (*i.e.*, cleavage or probe degradation due to sources other than the nucleases being compared). A good negative control colony would

be one containing only the vector used in the cloning and mutagenesis, for example, colonies containing only the pTrc99A vector.

Two factors that may influence the number of colonies chosen from a specific mutagenesis reaction for the initial rapid screen are 1) total number of colonies obtained from the mutagenesis reaction, and 2) whether the mutagenesis reaction was site-specific or randomly distributed across a whole gene or a region of a gene. For example, if only 5-10 colonies are present on the plate, all colonies can easily be tested. If hundreds of colonies are present, a subset of these may be analyzed. Generally 10-20 colonies are tested from a site-specific mutagenesis reaction, while 80 to 100 or more colonies are routinely tested from a single random mutagenesis reaction.

Where indicated, the altered 5' nucleases described in these experimental examples were tested as detailed below.

A. Rapid screen: INVADER activity on RNA target (Figure 18A)

A 2X substrate mix was prepared, comprising 20 mM MOPS, pH 7.5, 10 mM MgSO₄, 200 mM KCl, 2 μM FRET-probe oligo SEQ ID NO:223 (5'-FI-CGCT-cy3-TCTCGCTCGC-3'), 1 μM INVADER oligo SEQ ID NO:224 (5'-ACGGAACGAGCGTCTTTG-3'), and 4 nM RNA target SEQ ID NO:225 (5'-GCG AGC GAGA CAG CGA AAG ACG CUC GUU CCG U-3'). Five μl of the 2X substrate mix were dispensed into each sample well of a 96 well microtiter plate (Low Profile MULTIPLATE 96, M.J. Research, Inc.).

Cell suspensions were prepared by picking single colonies (mutants, positive control, and negative control colonies) and suspending each one in 20 μl of water. This can be done conveniently in a 96 well microtiter plate format, using one well per colony.

Five μl of the cell suspension was added to the appropriate test well such that the final reaction conditions were 10 mM MOPS, pH 7.5, 5 mM MgSO₄, 100 mM KCl, 1 μM FRET-probe oligo, 0.5 μM INVADER oligo, and 2 nM RNA target. The wells were covered with 10 μl of Clear CHILLOUT 14 (M.J. Research, Inc.) liquid wax, and the samples were heated at 85°C for 3 minutes, then incubated at 59°C for 1 hour. After the incubation, the plates were read on a Cytofluor fluorescence plate reader using the following parameters: excitation 485/20, emission 530/30.

B. Rapid screen: INVADER activity on DNA target (Figure 18B)

A 2X substrate mix was prepared, comprising 20 mM MOPS, pH 7.5, 10 mM MgSO₄, 200 mM KCl, 2 μM FRET-probe oligo SEQ ID NO:223

(5'-Fl-CGCT-Cy3-TCTCGCTCGC-3'), 1 μM INVADER oligo SEQ ID NO:224 (5'-ACGGAACGAGCGTCTTTG-3'), 1 nM DNA target SEQ ID NO:226 (5'-GCG AGC GAGA CAG CGA AAG ACG CTC GTT CCG T-3'). Five μl of the 2X substrate mix was dispensed into each sample well of a 96 well microtiter plate (MJ Low Profile).

Cell suspensions were prepared by picking single colonies (mutants, positive control and negative control colonies) and suspending them in 20 μl of water, generally in a 96 well microtiter plate format.

5 μl of the cell suspension were added to the appropriate test well such that the final reaction conditions were 10 mM MOPS, pH 7.5, 5 mM MgSO₄, 100 mM KCl, 1 μM FRET-probe oligo, 0.5 μM INVADER oligo, and 0.5 nM DNA target. Wells were covered with 10 μl of Clear CHILLOUT 14 (M.J. Research, Inc.) liquid wax, and the reactions were heated at 85°C for 3 minutes, then incubated at 59°C for 1 hour. After the hour incubation, the plate were read on a Cytofluor fluorescence plate reader using the following parameters: excitation 485/20, emission 530/30, gain 40, reads per well 10.

C. Rapid protein extraction (crude cell lysate)

Those mutants that gave a positive or an unexpected result in either the RNA or DNA INVADER assay were further analyzed, specifically for background activity on the X-structure or the hairpin substrate (Figure 18C and D, respectively). A rapid colony screen format can be employed, as described above. By simply changing the substrate, tests for background or aberrant enzymatic activity can be done. Another approach would be to do a rapid protein extraction from a small overnight culture of positive clones, and then test this crude cell lysate for additional protein function. One possible rapid protein extraction procedure is detailed below. Two to five ml of LB (containing the appropriate antibiotic for plasmid selection; *See e.g.*, Maniatis, books 1,2 and 3) were inoculated with the remaining volume of the 20 μl water-cell suspension and incubated at 37°C overnight. About 1.4 ml of the culture were transferred to a 1.5 ml microcentrifuge

tube, and microcentrifuged at top speed (e.g., 14,000 rpm in an Eppendorf 5417 table top microcentrifuge), at room temperature for 3-5 minutes to pellet the cells. The supernatant was removed, and the cell pellet was suspended in 100 μ l of TES buffer pH 7.5 (Sigma). Lysozyme (Promega) was added to a final concentration of 0.5 μ g/ μ l and samples were incubated at room temperature for 30 minutes. Samples were then heated at 70°C for 10 minutes to inactivate the lysozyme, and the cell debris was pelleted by microcentrifugation at top speed for 5 minutes. The supernatant was removed and this crude cell lysate was used in the following enzymatic activity assays.

D. Rapid screen: background specificity X structure substrate (Figure 18C)

Reactions were performed under conditions as detailed above. One μ l of crude cell lysate was added to 9 μ l of reaction components for a final volume of 10 μ l and final concentrations of 10 mM MOPS, pH 7.5, 5 mM MgSO₄, 100 mM KCl, 1 μ M FRET-probe oligo (SEQ ID NO:223), 0.5 μ M X-structure INVADER oligo SEQ ID NO:227 (5'-ACGGAACGAGCGTCTTTCATCTGTCAATC-3'), and 0.5 nM DNA target (SEQ ID NO:226). Wells were covered with 10 μ l of Clear CHILLOUT 14 (M.J. Research, Inc.) liquid wax, and the reactions were heated at 85°C for 3 minutes, then incubated at 59°C for 1 hour. After the incubation, the plates were read on a Cytofluor fluorescence plate reader using the following parameters: excitation 485/20, emission 530/30, gain 40, reads per well 10.

E. Rapid screen: background specificity hairpin substrate (Figure 18D)

Reactions were performed under conditions as detailed above. One μ l of crude cell lysate was added to 9 μ l of reaction components for a final volume of 10 μ l and final concentrations of 10 mM MOPS, pH 7.5, 5 mM MgSO₄, 100 mM KCl, 1 μ M FRET-probe oligonucleotide (SEQ ID NO:223), and 0.5 nM DNA target (SEQ ID NO:226). Wells were covered with 10 μ l of Clear CHILLOUT 14 (M.J. Research, Inc.) liquid wax, and the reactions were heated at 85°C for 3 minutes, then incubated at 59°C for 1 hour. After the hour incubation, the plate were read on a Cytofluor plate reader

using the following parameters: excitation 485/20, emission 530/30, gain 40, reads per well 10.

F. Activity assays with IrT1 and IdT targets (Figures 24)

The 5' nuclease activities assays were carried out in 10 μ l of a reaction containing 10 mM MOPS, pH 7.5, 0.05% Tween 20, 0.05% Nonidet P-40, 10 μ g/ml tRNA, 100 mM KCl and 5 mM MgSO₄. The probe concentration (SEQ ID NO: 167) was 2 mM. The substrates (IrT1 (SEQ ID NO: 228) or IdT (SEQ ID NO: 229) at 10 or 1 nM final concentration respectively) and approximately 20 ng of an enzyme, prepared as in Example 3, were mixed with the above reaction buffer and overlaid with CHILLOUT (MJ Research) liquid wax. Reactions were brought up to reaction temperature 57 °C, started by addition of MgSO₄, and incubated for 10 min. Reactions were then stopped by the addition of 10 μ l of 95% formamide containing 10 mM EDTA and 0.02% methyl violet (Sigma). Samples were heated to 90°C for 1 minute immediately before electrophoresis through a 20% denaturing acrylamide gel (19:1 cross-linked), with 7 M urea, and in a buffer of 45 mM Tris-borate, pH 8.3, 1.4 mM EDTA. Unless otherwise indicated, 1 μ l of each stopped reaction was loaded per lane. Gels were then scanned on an FMBIO-100 fluorescent gel scanner (Hitachi) using a 505 nm filter. The fraction of cleaved product was determined from intensities of bands corresponding to uncut and cut substrate with FMBIO Analysis software (version 6.0, Hitachi). The fraction of cleavage product did not exceed 20% to ensure that measurements approximated initial cleavage rates. The turnover rate was defined as the number of cleaved signal probes generated per target molecule per minute under these reaction conditions (1/min).

G. Activity assays with X structure (X) and hairpin (HP) targets (Figures 22)

The 5' nuclease activity assays were carried out in 10 μ l of a reaction containing 10 mM MOPS, pH 7.5, 0.05% Tween 20, 0.05% Nonidet P-40, 10 μ g/ml tRNA, 100 mM KCl and 5 mM MgSO₄. Each oligo for formation of either the hairpin structure assembly (22A, SEQ ID NOS: 230 and 231) assembly or the X structure assembly (22B, SEQ ID NOS: 230-232) was added to a final concentration of 1 μ M, and approximately 20 ng of test enzyme prepared as described in Example 3, were mixed with the above reaction

buffer and overlaid with CHILLOUT (MJ Research) liquid wax. Reactions were brought up to reaction temperature 60 °C, started by addition of MgSO₄, and incubated for 10 min. Reactions were then stopped by the addition of 10 µl of 95% formamide containing 10 mM EDTA and 0.02% methyl violet (Sigma). Samples were heated to 90°C for 1 minute immediately before electrophoresis through a 20% denaturing acrylamide gel (19:1 cross-linked), with 7 M urea, and in a buffer of 45 mM Tris-borate, pH 8.3, 1.4 mM EDTA. Unless otherwise indicated, 1 µl of each stopped reaction was loaded per lane. Gels were then scanned on an FMBIO-100 fluorescent gel scanner (Hitachi) using a 505 nm filter. The fraction of cleaved product was determined from intensities of bands corresponding to uncut and cut substrate with FMBIO Analysis software (version 6.0, Hitachi). The fraction of cleavage product did not exceed 20% to ensure that measurements approximated initial cleavage rates. The turnover rate was defined as the number of cleaved signal probes generated per target molecule per minute under these reaction conditions (1/min).

H. Activity assays with human IL-6 target (Figure 10)

The 5' nuclease activities assays were carried out in 10 µl reactions containing 10 mM MOPS, pH 7.5, 0.05% Tween 20, 0.05% Nonidet P-40, 10 µg/ml tRNA, 100 mM KCl and 5 mM MgSO₄. Reactions comprising the DNA IL-6 substrate contained 0.05 nM IL-6 DNA target (SEQ ID NO: 163) and 1 µM of each probe (SEQ ID NO: 162) and INVADER (SEQ ID NO: 161) oligonucleotides, and were carried out at 60°C for 30 min. Reactions comprising the IL-6 RNA target (SEQ ID NO: 160) were performed under the same conditions, except that the IL-6 RNA target concentration was 1 nM and the reactions were performed at 57°C for 60 min. Each reaction contained approximately 20 ng of test enzyme, prepared as described in Example 3.

I. Activity assays with synthetic r25mer target (Figure 23)

Reactions comprising the synthetic r25mer target (SEQ ID NO: 233) were carried out under the same reaction conditions (10 mM MOPS, pH 7.5, 0.05% Tween 20, 0.05% Nonidet P-40, 10 µg/ml tRNA, 100 mM KCl and 5 mM MgSO₄) and 1 µM of each probe (SEQ ID NO: 234) and INVADER (SEQ ID NO: 235) oligonucleotides, except that the

r25mer target concentration was 5 nM and the reactions were performed at 58°C for 60 min. Approximately 20 ng of each test enzyme was added to the reactions. Enzymes were prepared as described in Example 3.

Any of the tests described above can be modified to derive the optimal conditions for enzymatic activity. For example, enzyme titrations can be done to determine the optimal enzyme concentration for maximum cleavage activity, and lowest background signal. By way of example, but not by way of limitation, many of the mutant enzymes were tested at 10, 20 and 40 ng amounts. Similarly, a temperature titration can also be incorporated into the tests. Since modifying the structure of a protein can alter its temperature requirements, a range of temperatures can be tested to identify the condition best suited for the mutant in question.

Examples of the results from such screens (using approximately 20 ng of the mutant enzyme) are shown in Tables 3-8, and Figures 12, 14, 15, 19, and 25.

EXAMPLE 2

Cloning and Expression of 5' nucleases of DNA polymerases and mutant polymerases

A. DNA polymerases of *Thermus aquaticus* and *Thermus thermophilus*

1. Cloning of TaqPol and TthPol

Type A DNA polymerases from eubacteria of the genus *Thermus* share extensive protein sequence identity (90% in the polymerization domain, using the Lipman-Pearson method in the DNA analysis software from DNASTar, WI) and behave similarly in both polymerization and nuclease assays. Therefore, the genes for the DNA polymerase of *Thermus aquaticus* (TaqPol), *Thermus thermophilus* (TthPol) and *Thermus scotoductus* were used as representatives of this class. Polymerase genes from other eubacterial organisms, including, but not limited to, *Escherichia coli*, *Streptococcus pneumoniae*, *Mycobacterium smegmatis*, *Thermus thermophilus*, *Thermus sp.*, *Thermotoga maritima*, *Thermosipho africanus*, and *Bacillus stearothermophilus* are equally suitable.

a. Initial TaqPol Isolation: mutant TaqA/G

The *Taq* DNA polymerase gene was amplified by polymerase chain reaction from genomic DNA from *Thermus aquaticus*, strain YT-1 (Lawyer *et al.*, *supra*), using as
5 primers the oligonucleotides described in SEQ ID NOS:236 and 237. The resulting fragment of DNA has a recognition sequence for the restriction endonuclease *EcoRI* at the 5' end of the coding sequence and a *Bgl*II sequence at the 3' end of the coding strand. Cleavage with *Bgl*II leaves a 5' overhang or "sticky end" that is compatible with the end generated by *Bam*HI. The PCR-amplified DNA was digested with *EcoRI* and *Bam*HI.
10 The 2512 bp fragment containing the coding region for the polymerase gene was gel purified and then ligated into a plasmid that contains an inducible promoter.

In one embodiment of the invention, the pTTQ18 vector, which contains the hybrid *trp-lac (tac)* promoter, was used (M.J.R. Stark, *Gene* 5:255 [1987]). The *tac* promoter is under the control of the *E. coli lac* repressor protein. Repression allows the
15 synthesis of the gene product to be suppressed until the desired level of bacterial growth has been achieved, at which point repression is removed by addition of a specific inducer, isopropyl-b-D-thiogalactopyranoside (IPTG). Such a system allows the controlled expression of foreign proteins that may slow or prevent growth of transformants.

Particularly strong bacterial promoters, such as the synthetic *Ptac*, may not be
20 adequately suppressed when present on a multiple copy plasmid. If a highly toxic protein is placed under control of such a promoter, the small amount of expression leaking through, even in the absence of an inducer, can be harmful to the bacteria. In another embodiment of the invention, another option for repressing synthesis of a cloned gene product is contemplated. A non-bacterial promoter from bacteriophage T7, found in the
25 plasmid vector series pET-3, was used to express the cloned mutant *Taq* polymerase genes (Studier and Moffatt, *J. Mol. Biol.*, 189:113 [1986]). This promoter initiates transcription only by T7 RNA polymerase. In a suitable strain, such as BL21(DE3)pLYS, the gene for the phage T7 RNA polymerase is carried on the bacterial genome under control of the *lac* operator. This arrangement has the advantage that
30 expression of the multiple copy gene (on the plasmid) is completely dependent on the

expression of T7 RNA polymerase, which is easily suppressed because it is present in a single copy.

These are just two examples of vectors having suitable inducible promoters. Others are well known to those skilled in the art, and it is not intended that the improved nucleases of the present invention be limited by the choice of expression system.

For ligation into the pTTQ18 vector, the PCR product DNA containing the *Taq* polymerase coding region (termed mut*Taq* for reasons discussed below, SEQ ID NO:238) was digested with EcoRI and BglII and this fragment was ligated under standard "sticky end" conditions (Sambrook *et al. Molecular Cloning*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, pp. 1.63-1.69 [1989]) into the EcoRI and BamHI sites of the plasmid vector pTTQ18. Expression of this construct yields a translational fusion product in which the first two residues of the native protein (Met-Arg) are replaced by three from the vector (Met-Asn-Ser), but the remainder of the PCR product's protein sequence is not changed (SEQ ID NO:239). The construct was transformed into the JM109 strain of *E. coli*, and the transformants were plated under incompletely repressing conditions that do not permit growth of bacteria expressing the native protein. These plating conditions allow the isolation of genes containing pre-existing mutations, such as those that result from the infidelity of *Taq* polymerase during the amplification process.

Using this amplification/selection protocol, a clone was isolated containing a mutated *Taq* polymerase gene (mut*Taq*). The mutant was first detected by its phenotype, in which temperature-stable 5' nuclease activity in a crude cell extract was normal, but polymerization activity was almost absent (approximately less than 1% of wild type *Taq* polymerase activity). Polymerase activity was determined by primer extension reactions. The reactions were carried out in 10 µl of buffer containing 10 mM MOPS, pH 7.5, 5 mM MgSO₄, 100 mM KCl. In each reaction, 40 ng of enzyme were used to extend 10 µM (dT)₂₅₋₃₀ primer in the presence of either 10 µM poly (A)₂₈₆ or 1 µM poly (dA)₂₇₃ template, 45 µM dTTP and 5 µM Fl-dUTP at 60°C for 30 minutes. Reactions were stopped with 10 µl of stop solution (95% formamide, 10 mM EDTA, 0.02% methyl violet dye). Samples (3 µl) were fractionated on a 15% denaturing acrylamide gel (19:1

crossed-linked) and the fraction of incorporated Fl-dUTP was quantitated using an FMBIO-100 fluorescence gel scanner (Hitachi) equipped with a 505 nm emission filter.

DNA sequence analysis of the recombinant gene showed that it had changes in the polymerase domain resulting in two amino acid substitutions: an A to G change at nucleotide position 1394, which causes a Glu to Gly change at amino acid position 465 (numbered according to the natural nucleic and amino acid sequences, SEQ ID NOS:153 and 157), and another A to G change at nucleotide position 2260, which causes a Gln to Arg change at amino acid position 754. Because the Gln to Gly mutation is at a nonconserved position and because the Glu to Arg mutation alters an amino acid that is conserved in virtually all of the known Type A polymerases, the latter mutation is most likely the one responsible for curtailing the synthesis activity of this protein. The nucleotide sequence for the construct is given in SEQ ID NO:39. The enzyme encoded by this sequence is referred to as Taq A/G.

b. Initial TthPol Isolation

The DNA polymerase enzyme from the bacterial species *Thermus thermophilus* (Tth) was produced by cloning the gene for this protein into an expression vector and overproducing it in *E. coli* cells. Genomic DNA was prepared from 1 vial of dried *Thermus thermophilus* strain HB-8 from ATCC (ATCC #27634). The DNA polymerase gene was amplified by PCR using the following primers: 5'-CACGAATTCCGAGGCGATGCTTCGCTC-3' (SEQ ID NO:240) and 5'-TCGACGTCGACTAACCCTTGCGGAAAGCC-3' (SEQ ID NO:241). The resulting PCR product was digested with *EcoRI* and *SalI* restriction endonucleases and inserted into *EcoRI/Sal I* digested plasmid vector pTrc99G (described in Example 2C1) to create the plasmid pTrcTth-1. This *Tth* polymerase construct is missing a single nucleotide that was inadvertently omitted from the 5' oligonucleotide, resulting in the polymerase gene being out of frame. This mistake was corrected by site specific mutagenesis of pTrcTth-1 as described in Examples 4 and 5 using the following oligonucleotide: 5'-GCATCGCCTCGGAATTCATGGTC-3' (SEQ ID NO:242), to create the plasmid pTrcTth-2. The protein and the nucleic acid sequence encoding the protein are referred to as TthPol, and are listed as SEQ ID NOS:243 and 244 respectively.

c. Large Scale preparation of recombinant proteins

The recombinant proteins were purified by the following technique which is derived from a *Taq* DNA polymerase preparation protocol (Engelke *et al.*, Anal. Biochem., 191:396 [1990]) as follows. *E. coli* cells (strain JM109) containing either pTrc99A *Taq*Pol, pTrc99GTthPol were inoculated into 3 ml of LB containing 100 mg/ml ampicillin and grown for 16 hrs at 37°C. The entire overnight culture was inoculated into 200 ml or 350 ml of LB containing 100 mg/ml ampicillin and grown at 37°C with vigorous shaking to an A₆₀₀ of 0.8. IPTG (1 M stock solution) was added to a final concentration of 1 mM and growth was continued for 16 hrs at 37°C.

The induced cells were pelleted and the cell pellet was weighed. An equal volume of 2X DG buffer (100 mM Tris-HCl, pH 7.6, 0.1 mM EDTA) was added and the pellet was suspended by agitation. Fifty mg/ml lysozyme (Sigma) were added to 1 mg/ml final concentration and the cells incubated at room temperature for 15 min. Deoxycholic acid (10% solution) was added dropwise to a final concentration of 0.2 % while vortexing. One volume of H₂O and 1 volume of 2X DG buffer were added, and the resulting mixture was sonicated for 2 minutes on ice to reduce the viscosity of the mixture. After sonication, 3 M (NH₄)₂SO₄ was added to a final concentration of 0.2 M, and the lysate was centrifuged at 14000 x g for 20 min at 4°C. The supernatant was removed and incubated at 70°C for 60 min at which time 10% polyethylimine (PEI) was added to 0.25%. After incubation on ice for 30 min., the mixture was centrifuged at 14,000 x g for 20 min at 4°C. At this point, the supernatant was removed and the protein precipitated by the addition of (NH₄)₂SO₄ as follows.

Two volumes of 3 M (NH₄)₂SO₄ were added to precipitate the protein. The mixture was incubated overnight at room temperature for 16 hrs centrifuged at 14,000 x g for 20 min at 4°C. The protein pellet was suspended in 0.5 ml of Q buffer (50 mM Tris-HCl, pH 8.0, 0.1 mM EDTA, 0.1% Tween 20). For the *Mja* FEN-1 preparation, solid (NH₄)₂SO₄ was added to a final concentration of 3 M (~75% saturated), the mixture was incubated on ice for 30 min, and the protein was spun down and suspended as described above.

The suspended protein preparations were quantitated by determination of the A₂₇₉ dialyzed and stored in 50% glycerol, 20 mM Tris HCl, pH8.0, 50 mM KCl, 0.5% Tween 20, 0.5% Nonidet P-40, with 100 µg/ml BSA.

B. DNA polymerases of *Thermus filiformis* and *Thermus scotoductus*

1. Cloning of *Thermus filiformis* and *Thermus scotoductus*

One vial of lyophilized *Thermus filiformis* (Tfi) obtained from DSMZ (Deutsche Sammlung von Mikroorganismen und Zellculturen, Braunschweig, Germany, strain #4687) was rehydrated in 1 ml of Castenholz medium (DSMZ medium 86) and inoculated into 500 ml of Castenholz medium preheated to 50°C. The culture was incubated at 70°C with vigorous shaking for 48 hours. After growth, the cells were harvested by centrifugation at 8000 x g for 10 minutes, the cell pellet was suspended in 10 ml of TE (10 mM TrisHCL, pH 8.0, 1 mM EDTA), and the cells were frozen at -20°C in 1 ml aliquots. A 1 ml aliquot was thawed, lysozyme was added to 1 mg/ml, and the cells were incubated at 23°C for 30 minutes. A solution of 20% SDS (sodium dodecyl sulfate) was added to a final concentration of 0.5% followed by extraction with buffered phenol. The aqueous phase was further extracted with 1:1 phenol:chloroform, and extracted a final time with chloroform. One-tenth volume of 3 M sodium acetate, pH 5.0 and 2.5 volumes of ethanol were added to the aqueous phase and mixed. The DNA was pelleted by centrifugation at 20,000 x g for 5 minutes. The DNA pellet was washed with 70% ethanol, air dried and resuspended in 200 µl of TE and used directly for amplification. *Thermus scotoductus* (Tsc, ATCC # 51532) was grown and genomic DNA was prepared as described above for *Thermus filiformis*.

The DNA polymerase I gene from Tfi (GenBank accession #AF030320) could not be amplified as a single fragment. Therefore, it was cloned in 2 separate fragments into the expression vector pTrc99a. The 2 fragments overlap and share a *Not I* site which was created by introducing a silent mutation at position 1308 of the Tfi DNA polymerase open reading frame (ORF) in the PCR oligonucleotides. The 3' half of the gene was amplified using the Advantage cDNA PCR kit (Clontech) with the following oligonucleotides; 5'-ATAGCCATGGTGGAGCGGCCGCTCTCCCGG (SEQ ID NO:245) and 5'-AAGCGTCGACTCAATCCTGCTTCGCCTCCAGCC (SEQ ID NO:246). The PCR

product from this reaction was approximately 1200 base pairs in length. It was cut with the restriction enzymes Not I and Sal I, and the resulting DNA was ligated into pTrc99a cut with NotI and SalI to create pTrc99a-Tfi3'. The 5' half of the gene was amplified as described above using the following two primers;

5 5'AATCGAATTACCCCCACTTTTGTGACCTGGAGG (SEQ ID NO:247) and
5'-CCGGGAGAGCGGCCGCTCCAC (SEQ ID NO:248). The resulting 1300 base pair fragment was cut with restriction enzymes Eco RI and Not I and ligated into pTrc99a-Tfi3' cut with *NotI* and *EcoRI* to produce pTrc99a-TfiPol, SEQ ID NO:249 (the corresponding amino acid sequence is listed in SEQ ID NO:250).

10 The DNA polymerase I gene from *Thermus scotoductus* was amplified using the Advantage cDNA PCR kit (Clontech) using the following two primers;

5'-ACTGGAATTCCTGCCCCTCTTTGAGCCCAAG (SEQ ID NO:251) and
5'-AACAGTCGACCTAGGCCTTGGCGGAAAGCC (SEQ ID NO:252). The PCR product was cut with restriction enzymes Eco RI and Sal I and ligated into Eco RI, Sal I
15 cut pTrc99a to create pTrc99a-TscPol SEQ ID NO:253 (the corresponding amino acid sequence is listed in SEQ ID NO:254).

2. Expression and purification of *Thermus filiformis* and *Thermus scotoductus*

20 Plasmids were transformed into protease deficient *E. coli* strain BL21 (Novagen) or strain JM109 (Promega Corp., Madison, WI) for protein expression. Flasks containing 200 ml of LB containing 100 µg/ml ampicillin were inoculated with either a single colony from an LB plate or from a frozen stock of the appropriate strain. After several hours of growth at 37°C with vigorous shaking, cultures was induced by the addition of
25 200 µl of 1 M isothiopyl-galatoside (IPTG). Growth at 37°C was continued for 16 hours prior to harvest. Cells were pelleted by centrifugation at 8000 x g for 15 minutes followed by suspension of the cell pellet in 5 ml of TEN (10 mM TrisHCl, pH 8.0, 1 mM EDTA, 100 mM NaCl). 100 µl of 50 mg/ml lysozyme were added and the cells incubated at room temperature for 15 minutes. Deoxycholic acid (10%) was added to a
30 final concentration of 0.2%. After thorough mixing, the cell lysates were sonicated for 2 minutes on ice to reduce the viscosity of the mixture. Cellular debris was pelleted by

centrifugation at 4°C for 15 minutes at 20,000 x g. The supernatant was removed and incubated at 70°C for 30 min after which 10% polyethylimine (PEI) was added to 0.25%. After incubation on ice for 30 minutes, the mixture was centrifuged at 20,000 x g for 20 min at 4°C. At this point, the supernatant containing the enzyme was removed, and the protein was precipitated by the addition of 1.2 g of ammonium sulfate and incubation at 4°C for 1 hour. The protein was pelleted by centrifugation at 4°C for 10 minutes at 20,000 x g. The pellet was resuspended in 4 ml of HPLC Buffer A (50 mM TrisHCl, pH 8.0, 1 mM EDTA). The protein was further purified by affinity chromatography using an Econo-Pac heparin cartridge (Bio-Rad) and a Dionex DX 500 HPLC instrument. Briefly, the cartridge was equilibrated with HPLC Buffer A, and the enzyme extract was loaded on the column and eluted with a linear gradient of NaCl (0-2 M) in the same buffer. Pure protein elutes between 0.5 and 1 M NaCl. The enzyme peak was collected and dialyzed in 50% glycerol, 20 mM Tris HCl, pH 8, 50 mM KCl, 0.5% Tween 20, 0.5% Nonidet P40, 100 mg/ml BSA.

C. Generation of polymerase mutants with reduced polymerase activity but unaltered 5' nuclease activity

All mutants generated in section C were expressed and purified as described in Example 2A1C.

1. Modified TaqPol Genes: TaqDN

A polymerization deficient mutant of *Taq* DNA polymerase called TaqDN was constructed. TaqDN nuclease contains an asparagine residue in place of the wild-type aspartic acid residue at position 785 (D785N).

DNA encoding the TaqDN nuclease was constructed from the gene encoding the Taq A/G in two rounds of site-directed mutagenesis. First, the G at position 1397 and the G at position 2264 of the Taq A/G gene (SEQ ID NO:238) were changed to A at each position to recreate a wild-type TaqPol gene. In a second round of mutagenesis, the wild type TaqPol gene was converted to the Taq DN gene by changing the G at position 2356 to A. These manipulations were performed as follows.

DNA encoding the Taq A/G nuclease was recloned from pTTQ18 plasmid into the pTrc99A plasmid (Pharmacia) in a two-step procedure. First, the pTrc99A vector was modified by removing the G at position 270 of the pTrc99A map, creating the pTrc99G cloning vector. To this end, pTrc99A plasmid DNA was cut with *NcoI* and the recessive 3' ends were filled-in using the Klenow fragment of *E.coli* polymerase I in the presence of all four dNTPs at 37°C for 15 min. After inactivation of the Klenow fragment by incubation at 65°C for 10 min, the plasmid DNA was cut with *EcoRI* and the ends were again filled-in using the Klenow fragment in the presence of all four dNTPs at 37°C for 15 min. The Klenow fragment was then inactivated by incubation at 65°C for 10 min. The plasmid DNA was ethanol precipitated, recircularized by ligation, and used to transform *E.coli* JM109 cells (Promega). Plasmid DNA was isolated from single colonies, and deletion of the G at position 270 of the pTrc99A map was confirmed by DNA sequencing.

In a second step, DNA encoding the Taq A/G nuclease was removed from the pTTQ18 plasmid using *EcoRI* and *SalI* and the DNA fragment carrying the Taq A/G nuclease gene was separated on a 1% agarose gel and isolated with GeneClean II Kit (Bio 101, Vista, CA). The purified fragment was ligated into the pTrc99G vector that had been cut with *EcoRI* and *SalI*. The ligation mixture was used to transform competent *E.coli* JM109 cells (Promega). Plasmid DNA was isolated from single colonies and insertion of the Taq A/G nuclease gene was confirmed by restriction analysis using *EcoRI* and *SalI*.

Plasmid DNA pTrcAG carrying the Taq A/G nuclease gene cloned into the pTrc99A vector was purified from 200 ml of JM109 overnight culture using QIAGEN Plasmid Maxi kit (QIAGEN, Chatsworth, CA) according to manufacturer's protocol. pTrcAG plasmid DNA was mutagenized using two mutagenic primers, E465 (SEQ ID NO:255) (Integrated DNA Technologies, Iowa) and R754Q (SEQ ID NO:256) (Integrated DNA Technologies), and the selection primer Trans Oligonucleotide AlwNI/SpeI (Clontech, Palo Alto, CA, catalog #6488-1) according to TRANSFORMER Site-Directed Mutagenesis Kit protocol (Clontech, Palo Alto, CA) to produce a restored wild-type TaqPol gene (pTrcWT).

The resulting mutant genes were termed Taq DN HT (SEQ ID NO:265, nucleic acid sequence; SEQ ID NO:266, amino acid sequence) and Tth DN HT (SEQ ID NO:267, nucleic acid sequence; SEQ ID NO:268, amino acid sequence).

4. Purification of Taq DN HT and Tth DN HT

Both Taq DN HT and Tth DN HT proteins were expressed in *E. coli* strain JM109 as described in Example 2B2. After ammonium sulfate precipitation and centrifugation, the protein pellet was suspended in 0.5 ml of Q buffer (50 mM Tris-HCl, pH 8.0, 0.1 mM EDTA, 0.1% Tween 20). The protein was further purified by affinity chromatography using His-Bind Resin and Buffer Kit (Novagen) according to the manufacturer's instructions. 1 ml of His-Bind resin was transferred into a column, washed with 3 column volumes of sterile water, charged with 5 volumes of 1X Charge Buffer, and equilibrated with 3 volumes of 1X Binding Buffer. Four ml of 1X Binding Buffer was added to the protein sample and the sample solution was loaded onto the column. After washing with 3 ml of 1X Binding Buffer and 3 ml of 1X Wash Buffer, the bound His-Tag protein was eluted with 1 ml of 1X Elute Buffer. The pure enzyme was then dialyzed in 50% glycerol, 20 mM Tris-HCl, pH 8.0, 50 mM KCl, 0.5% Tween 20, 0.5% Nonidet P40, and 100 µg/ml BSA. Enzyme concentrations were determined by measuring absorption at 279 nm.

EXAMPLE 3

RNA-dependent 5' nuclease activity of TthPol can be conferred on TaqPol by transfer of the N-terminal portion of the DNA polymerase domain

A. Preparation and purification of substrate structures having either a DNA or an RNA target strand

The downstream (SEQ ID NO:162) and upstream probes (SEQ ID NO:161) and the IL-6 DNA (SEQ ID NO:163) (Figure 10) target strand were synthesized on a PerSeptive Biosystems instrument using standard phosphoramidite chemistry (Glen Research). The synthetic RNA-DNA chimeric IrT target labeled with biotin at the 5'-end

5 (Figure 20A) was synthesized utilizing 2'-ACE RNA chemistry (Dharmacon Research). The 2'-protecting groups were removed by acid-catalyzed hydrolysis according to the manufacturer's instructions. The downstream probes labeled with 5'-fluorescein (Fl) or 5'-tetrachloro-fluorescein (TET) at their 5' ends were purified by reverse phase HPLC using a Resource Q column (Amersham-Pharmacia Biotech). The 648-nucleotide IL-6 RNA target (SEQ ID NO:160) (Figure 10) was synthesized by T7 RNA polymerase runoff-transcription of the cloned fragment of human IL-6 cDNA (nucleotides 64-691 of the sequence published in May *et al.*, Proc. Natl. Acad. Sci., 83:8957 [1986]) using a Megascript Kit (Ambion). All oligonucleotides were finally purified by separation on a 10 20% denaturing polyacrylamide gel followed by excision and elution of the major band. Oligonucleotide concentration was determined by measuring absorption at 260 nm. The biotin labeled IrT target was incubated with a 5-fold excess of streptavidin (Promega) in a buffer containing 10 mM MOPS, pH 7.5, 0.05% Tween 20, 0.05% NP-40 and 10 µg/ml tRNA at room temperature for 10 min.

15 **B. Introduction of restriction sites to make chimeras**

The restriction sites used for formation of chimerical proteins, described below, were chosen for convenience. The restriction sites in the following example have been strategically placed to surround regions shown by crystal structure and other analysis to be functional domains (*See*, Figures 6, 7, and 19). Different sites, either naturally occurring or created via directed mutagenesis can be used to make similar constructs with other Type A polymerase genes from related organisms. It is desirable that the mutations all be silent with respect to protein function. By studying the nucleic acid sequence and the amino acid sequence of the protein, one can introduce changes in the nucleic acid 25 sequence that have no effect on the corresponding amino acid sequence. If the nucleic acid change required affects an amino acid, one can make the alteration such that the new amino acid has the same or similar characteristics of the one replaced. If neither of these options is possible, one can test the mutant enzyme for function to determine if the nucleic acid alteration caused a change in protein activity, specificity or function. It is 30 not intended that the invention be limited by the particular restriction sites selected or introduced for the creation of the improved enzymes of the present invention.

C. Generation of Tth DN RX HT and Taq DN RX HT

Mutagenesis was performed to introduce 3 additional, unique restriction sites into the polymerase domain of both the Taq DN HT and Tth DN HT enzymes. Site-specific mutagenesis was performed using the Transformer Site-Directed Mutagenesis Kit from (Clontech) according to manufacturer's instructions. One of two different selection primers, Trans Oligo AlwNI/SpeI or Switch Oligo SpeI/AlwNI (Clontech, Palo Alto CA catalog #6488-1 or catalog #6373-1) was used for all mutagenesis reactions described. The selection oligo used in a given reaction is dependent on the selection restriction site present in the vector. All mutagenic primers were synthesized by standard synthetic chemistry. Resultant colonies were expressed in *E.coli* strain JM109.

The Not I sites (amino acid position 328) were created using the mutagenic primers 5'-gccgccagggcgccgcgtccaccgggcc (SEQ ID NO:269) and 5'-gcctgcagggcgccgcgtgcaccggggca (SEQ ID NO:270) corresponding to the sense strands of the Taq DN HT and the Tth DN HT genes, respectively. The BstI (amino acid position 382) and NdeI (amino acid position 443) sites were introduced into both genes using sense strand mutagenic primers 5'-ctcctggacccttcgaacaccacccc (SEQ ID NO:271) and 5'-gtcctggcccatatggaggccac (SEQ ID NO:272). The mutant plasmids were over-expressed and purified using Qiagen QiaPrep Spin Mini Prep Kit (cat. # 27106). The vectors were tested for the presence of the restriction sites by DNA sequencing and restriction mapping. These constructs are termed Tth DN RX HT (DNA sequence SEQ ID NO:273; amino acid sequence SEQ ID NO:274) and Taq DN RX HT (DNA sequence SEQ ID NO:275; amino acid sequence SEQ ID NO:276).

D. Chimeras

The chimeric constructs shown in Figure 19 were created by exchanging homologous DNA fragments defined by the restriction endonuclease sites EcoRI (E) and BamHI (B), common for both genes, the cloning vector site SalI (S) and the new sites, NotI (N), BstBI (Bs), NdeI (D) created at the homologous positions of both genes by site directed mutagenesis. In generating these chimeric enzymes, two different pieces of DNA are ligated together to yield the final construct. The larger piece of DNA that

contains the plasmid vector as well as part of the Taq or Tth (or parts of both) sequence will be termed the “vector.” The smaller piece of DNA that contains sequences of either the Taq or Tth (or parts of both) polymerase will be termed the “insert.”

All restriction enzymes were from New England Biolabs or Promega and used in reactions with the accompanying buffer, according to the manufacturer’s instructions.

Reactions were done in 20 µl volume with about 500 ng of DNA per reaction, at the optimal temperature for the specified enzyme. More than one enzyme was used in a single reaction (double digest) if the enzymes were compatible with respect to reaction buffer conditions and reaction temperature. If the enzymes in question were not compatible with respect to buffer conditions, the enzyme requiring the lowest salt condition was used first. After the completion of that reaction, buffer conditions were changed to be optimal or better suited to the second enzyme, and the second reaction was performed. These are common restriction enzyme digest strategies, well known to those in the art of basic molecular biology (Maniatis, supra).

The digested restriction fragments were gel isolated for optimal ligation efficiency. Two µl of 10X loading dye (50% glycerol, 1X TAE, 0.5% bromophenol blue) were added to the 20 µl reaction. The entire volume was loaded and run on a 1%, 1X TAE agarose gel containing 1 µl of a 1% ethidium bromide solution per 100 ml of agarose gel solution. The digested fragments were visualized under UV light, and the appropriate fragments (as determined by size) were excised from the gel. These fragments were then purified using the Qiagen Gel Extractio Kit, (cat # 28706) according to the manufacturer’s instructions.

Ligations were performed in a 10 µl volume, using 400 units per reaction of T4 DNA Ligase enzyme from New England Biolabs (catalog #202L), with the accompanying reaction buffer. Ligation reactions were done at room temperature for 1 hour, with 1 µl of each of the Qiagen-purified fragments (approximately 20-50 ng of each DNA, depending on recovery from the gel isolation). Ligation products were then transformed into *E. coli* strain JM 109 and plated onto an appropriate growth and selection medium, such as LB with 100µg/ml of ampicillin to select for transformants.

For each ligation reaction, six transformants were tested to determine if the desired construct was present. Plasmid DNA was purified and isolated using the QiaPrep

Spin Mini Prep Kit, according to manufacturer's instructions. The constructs were verified by DNA sequencing and by restriction mapping.

Expression and purification of the chimeric enzymes was done as follows.

Plasmids were transformed into *E. coli* strain JM109 (Promega). Log phase cultures (200 ml) of JM109 were induced with 0.5 mM IPTG (Promega) and grown for an additional 16 hours prior to harvest. Crude extracts containing soluble proteins were prepared by lysis of pelleted cells in 5 ml of 10 mM Tris-HCl, pH 8.3, 1mM EDTA, 0.5mg/ml lysozyme during incubation at room temperature for 15 minutes. The lysate was mixed with 5 ml of 10 mM Tris-HCl pH 7.8, 50 mM KCl, 1 mM EDTA, 0.5% Tween 20, 0.5% Nonidet P-40, heated at 72°C for 30 minutes, and cell debris was removed by centrifugation at 12,000x g for 5 minutes. Final purification of the protein was done by affinity chromatography using an Econo-Pac heparin cartridge (Bio-Rad) and Dionex DX 500 HPLC instrument. Briefly, the cartridge was equilibrated with 50mM Tris-HCl pH 8, 1 mM EDTA, and an enzyme extract dialyzed against the same buffer was loaded on the column and eluted with a linear gradient of NaCl (0-2 M) in the same buffer. The HPLC-purified protein was dialyzed and stored in 50% (vol/vol) glycerol, 20 mM Tris-HCl pH 8.0, 50 mM KCl, 0.5% Tween 20, 0.5% Nonidet P-40, and 100µg/m BSA. The enzymes were purified to homogeneity according to SDS-PAGE, and the enzyme concentrations were determined by measuring absorption at 279 nm.

1. Construction of TaqTth(N) and TthTaq(N)

The first exchange that was performed involved the polymerase domains of the two enzymes. Separation of the nuclease domain (the N-terminal end of the protein) from the polymerase domain (the C-terminal portion of the protein) was accomplished by cutting both genes with the restriction endonucleases EcoRI and NotI. The approximately 900 base pair fragment from the Tth DN RX HT gene was cloned into the homologous sites of the Taq DN RX HT gene, and the approximately 900 base pair fragment from the Taq DN RX HT gene was cloned into the homologous sites of the Tth DN RX HT gene, yielding two chimeras, TaqTth(N) (DNA sequence SEQ ID NO:69; amino acid sequence SEQ ID NO:2) which has the Taq DN RX HT 5' nuclease domain and the Tth DN RX HT polymerase domain, and TthTaq(N) (DNA sequence SEQ ID

NO:70; amino acid sequence SEQ ID NO:3) which is made up of the Tth DN RX HT 5' nuclease domain and the Taq DN RX HT polymerase domain.

2. Construction of TaqTth(N-B)

The Taq DN RX HT construct was cut with the enzymes NdeI and BamHI and the larger, vector fragment was gel isolated as detailed above. The Tth DN RX HT construct was also cut with NdeI and BamHI and the smaller (approximately 795 base pairs) Tth fragment was gel isolated and purified. The Tth NdeI-BamHI insert was ligated into the Taq NdeI-BamHI vector as detailed above to generate the TaqTth(N-B) (DNA sequence SEQ ID NO:71; amino acid sequence SEQ ID NO:4).

3. Construction of TaqTth(B-S)

The Taq DN RX HT construct was cut with the enzymes BamHI and SalI and the larger vector fragment was gel isolated as detailed above. The Tth DN RX HT construct was also cut with BamHI and SalI and the smaller (approximately 741 base pairs) Tth fragment was gel isolated and purified. The Tth BamHI-SalI insert was ligated into the Taq BamHI-SalI vector as detailed above to generate the TaqTth(B-S) (DNA sequence SEQ ID NO:72; amino acid sequence SEQ ID NO:5).

4. Construction of TaqTth(N-D)

The Taq DN RX HT construct was cut with the enzymes NotI and NdeI and the larger vector fragment was isolated as detailed above. The Tth DN RX HT construct was also cut with NotI and NdeI and the smaller (approximately 345 base pairs) Tth fragment was gel isolated and purified. The Tth NotI-NdeI insert was ligated into the Taq NotI-NdeI vector as detailed above to generate the TaqTth(N-D) (DNA sequence SEQ ID NO:73; amino acid sequence SEQ ID NO:6).

5. Construction of TaqTth(D-B)

The Taq DN RX HT construct was cut with the enzymes NdeI and BamHI and the larger vector fragment was isolated as detailed above. The Tth DN RX HT construct was also cut with NdeI and BamHI and the smaller (approximately 450 base pairs) Tth

fragment was gel isolated and purified. The Tth NdeI-BamHI insert was ligated into the Taq NdeI-BamHI vector as detailed above to generate the TaqTth(D-B) (DNA sequence SEQ ID NO:74; amino acid sequence SEQ ID NO:7).

6. Construction of TaqTth(Bs-B)

The Taq DN RX HT construct was cut with the enzymes BstBI and BamHI and the larger vector fragment was isolated as detailed above. The Tth DN RX HT construct was also cut with BstBI and BamHI and the smaller (approximately 633 base pairs) Tth fragment was gel isolated and purified. The Tth NdeI-BamHI insert was ligated into the Taq NdeI-BamHI vector as detailed above to generate TaqTth(Bs-B) (DNA sequence SEQ ID NO:75; amino acid sequence SEQ ID NO:8).

7. Construction of TaqTth(N-Bs)

The Taq DN RX HT construct was cut with the enzymes NotI and BstBI and the larger vector fragment was isolated as detailed above. The Tth DN RX HT construct was also cut with NotI and BstBI and the smaller (approximately 162 base pairs) Tth fragment was gel isolated and purified. The Tth NotI-BstBI insert was ligated into the Taq NotI-BstBI vector as detailed above to generate TaqTth(N-Bs) (DNA sequence SEQ ID NO:76; amino acid sequence SEQ ID NO:9).

8. Construction of TthTaq(B-S)

The Tth DN RX HT construct was cut with the enzymes BamHI and SalI and the larger vector fragment was isolated as detailed above. The Taq DN RX HT construct was also cut with BamHI and SalI and the smaller (approximately 741 base pairs) Tth fragment was gel isolated and purified. The Taq BamHI-SalI insert was ligated into the Tth BamHI-SalI vector as detailed above to generate the TthTaq(B-S) (DNA sequence SEQ ID NO:77; amino acid sequence SEQ ID NO:10).

9. Construction of Tth Taq(N-B)

The Tth DN RX HT construct was cut with the enzymes NotI and BamHI and the larger vector fragment was isolated as detailed above. The Taq DN RX HT construct was

also cut with NotI and BamHI and the smaller (approximately 795 base pairs) Tth fragment was gel isolated and purified. The Taq NotI-BamHI insert was ligated into the Tth NotI-BamHI vector as detailed above to generate the TthTaq(N-B) (DNA sequence SEQ ID NO:78; amino acid sequence SEQ ID NO:11).

5

The cleavage activities of these chimerical proteins were characterized as describe in Example 1, part A, and a comparison of the cleavage cycling rates on an RNA target is shown in Figure 12. As further discussed in the Description of the Invention, these data show that elements found in the central third of the TthPol protein are important in conferring the TthPol-like RNA-dependent cleavage activity on the chimerical proteins comprising portions of TaqPol.

10

EXAMPLE 4

Alterations influencing RNA-dependent 5' nuclease activity do not necessarily influence RNA-dependent DNA polymerase activity

15

TthPol is known to have a more active RNA template dependent DNA polymerase than does the TaqPol (Myers and Gelfand, Biochemistry 30:7661 [1991]). To determine whether the RNA template dependent 5' nuclease activity of the Thermus DNA Pol I enzymes is related to their RNA-dependent polymerase activity, the D785N and D787N mutations used to create the polymerase-deficient versions of TaqPol and TthPol, respectively were reversed. Polymerase activity was similarly restored to the TaqTth (N) (DNA sequence SEQ ID NO:79; amino acid sequence SEQ ID NO:12), TaqTth(N-B) (DNA sequence SEQ ID NO:80; amino acid sequence SEQ ID NO:13), TaqTth(B-S) (DNA sequence SEQ ID NO:81; amino acid sequence SEQ ID NO:14) chimeras, and the TaqPol(W417L/G418K/ E507Q) (DNA sequence SEQ ID NO:82; amino acid sequence SEQ ID NO:15) mutant proteins.

20

25

Polymerase function was restored in all the above mentioned enzyme mutants by inserting the BamHI to SalI fragment of the native, non-DN sequence into the selected chimera or mutant enzyme. For example, the mutant construct TaqTth(N-B) was cut with the restriction enzyme BamHI (approximate amino acid position 593) and the

30

restriction enzyme Sall (approximate amino acid position 840). The larger vector fragment was gel purified as described in Example 3D. The native TaqPol construct was also cut with the restriction endonucleases BamHI and Sall, and the smaller insert fragment containing the native amino acid sequence was also gel purified. The insert
5 fragment was then ligated into the vector as detailed in Experimental Example 3D.

The polymerase activities of these proteins were evaluated by extension of the dT₂₅₋₃₅-oligonucleotide primer with fluorescein-labeled dUTP in the presence of either poly(dA) or poly(A) template. Primer extension reactions were carried out in 10 µl buffer containing 10 mM MOPS, pH7.5, 5 mM MgSO₄, 100 mM KCl. Forty ng of
10 enzyme were used to extend 10 µM (dT)₂₅₋₃₀ primer in the presence of either 10 µM poly(A)₂₈₆ or 1 µM poly(dA)₂₇₃ template, 45 µM dTTP and 5 µM Fl-dUTP at 60°C for 30 min. Reactions were stopped with 10 µl of stop solution (95% formamide, 10 mM EDTA, 0.02% methyl violet dye). Samples (3 µl) were fractionated on a 15% denaturing acrylamide gel and the fraction of incorporated Fl-dUTP was quantitated using an
15 FMBIO-100 fluorescent gel scanner (Hitachi) equipped with a 505 nm filter as described above.

As shown in Figure 16, the DNA-dependent polymerase activities are very similar for all constructs used in this experiment, whereas the RNA-dependent polymerase activities of TthPol, TaqTth(N) and TaqTth(B-S) are at least 6-fold higher than the
20 activities of TaqPol, TaqTth(N-B) and the TaqPol W417L/G418K/E507Q mutant. From the analysis of these results, it can be concluded that the high RNA-dependent DNA polymerase activity of TthPol is determined by the C-terminal half of the polymerase domain (roughly, amino acids 593-830) and that the RNA-dependent 5' nuclease and polymerase activities are not related to each other, and are controlled by different regions.

EXAMPLE 5

Specific point mutants in Taq DN RX HT developed from information from the chimeric studies

The chimeric studies (Example 3, above) suggest that the part of the TthPol sequence determining its high RNA-dependent 5' nuclease activity comprises the

BstBI-BamHI region located approximately between amino acid 382 and 593.

Comparison of the amino acid sequences between the BstBI and BamHI regions of Tth DN RX HT and Taq DN RX HT (SEQ ID NOS:165 and 164, respectively) revealed only 25 differences (Figure 13). Among these, 12 amino acid changes were conservative while 13 of the differences resulted in a changes in charge. Since the analysis of the chimeric enzymes suggested that the critical mutations are located in both the BstBI-NdeI and the NdeI-BamHI regions of Tth DN RX HT, site specific mutagenesis was used to introduce the Tth DN RX HT specific amino acids into the BstBI-NdeI and NdeI-BamHI regions of the TaqTth(D-B) and the TaqTth(N-D) respectively.

Six Tth DN RX HT specific substitutions were generated in the BstBI-NdeI region of the TaqTth(D-B) by single or double amino acid mutagenesis. Similarly, 12 Tth DN RX HT specific amino acid changes were introduced at the homologous position of the NdeI-BamHI region of the TaqTth(N-D).

Plasmid DNA was purified from 200 ml of JM109 overnight culture using QIAGEN Plasmid Maxi Kit (QIAGEN, Chatsworth, CA) according to the manufacturer's protocol to obtain enough starting material for all mutagenesis reactions. All site specific mutations were introduced using the Transformer Site Directed mutagenesis Kit (Clontech) according to the manufacturer's protocol; specific sequence information for the mutagenic primers used for each site is provided below. One of two different selection primers, Trans Oligo AlwNI/SpeI or Switch Oligo SpeI/AlwNI (Clontech, Palo Alto, CA catalog #6488-1 or catalog #6373-1) was used for all mutagenesis reactions described. The selection oligo used in a given reaction is dependent on the restriction site present in the vector. All mutagenic primers were synthesized by standard synthetic chemistry. Resultant colonies were *E.coli* strain JM109.

1. **Construction of TaqTth(D-B) E404H (DNA sequence SEQ ID NO:83; amino acid sequence SEQ ID NO:16)**

Site specific mutagenesis was performed on pTrc99A TaqTth(D-B) DNA using the mutagenic primer 240-60-01 5'-gag gag gcg ggg cac cgg gcc gcc ctt-3' (SEQ ID NO:277) to introduce the E404H mutation.

2. Construction of TaqTth(D-B) F413H/A414R (DNA sequence SEQ ID NO:84; amino acid sequence SEQ ID NO:17)

Site specific mutagenesis was performed on pTrc99A TaqTth(D-B) DNA using the mutagenic primer 240-60-02 5'-ctt tcc gag agg ctc cat cgg aac ctg tgg ggg agg-3' (SEQ ID NO:278) to introduce the F413H and the A414R mutations.

3. Construction of TaqTth(D-B) W417L/G418K (DNA sequence SEQ ID NO:85; amino acid sequence SEQ ID NO:18)

Site specific mutagenesis was performed on pTrc99A TaqTth(D-B) DNA using the mutagenic primer 240-60-03 5'-ctc ttc gcc aac ctg ctt aag agg ctt gag ggg gag-3' (SEQ ID NO:279) to introduce the W417L and the G418K mutations.

4. Construction of TaqTth(D-B) A439R (DNA sequence SEQ ID NO:86; amino acid sequence SEQ ID NO:19)

Site specific mutagenesis was performed on pTrc99A TaqTth(ND-B) DNA using the mutagenic primer 240-60-04 5'-agg ccc ctt tcc cgg gtc ctg gcc cat-3' (SEQ ID NO:280) to introduce the A439R mutation.

5. Construction of TaqTth(N-D) L451R (DNA sequence SEQ ID NO:87; amino acid sequence SEQ ID NO:20)

Site specific mutagenesis was performed on pTrc99A TaqTth(N-D) DNA using the mutagenic primer 240-60-05 5'-acg ggg gtg cgc cgg gac gtg gcc tat-3' (SEQ ID NO:281) to introduce the L415 mutation.

6. Construction of TaqTth(N-D) R457Q (DNA sequence SEQ ID NO:88; amino acid sequence SEQ ID NO:21)

Site specific mutagenesis was performed on pTrc99A TaqTth(N-D) DNA using the mutagenic primer 240-60-06 5'-gtg gcc tat ctc cag gcc ttg tcc ctg-3' (SEQ ID NO:282) to introduce the L415Q mutation.

**7. Construction of TaqTth(N-D) V463L (DNA sequence SEQ ID NO:89;
amino acid sequence SEQ ID NO:22)**

Site specific mutagenesis was performed on pTrc99AtaqTth(N-D) DNA using the mutagenic primer 240-60-07 5'-ttg tcc ctg gag ctt gcc gag gag atc-3' (SEQ ID NO:283)

to introduce the V463L mutation.

**8. Construction of TaqTth(N-D) A468R (DNA sequence SEQ ID NO:90;
amino acid sequence SEQ ID NO:23)**

Site specific mutagenesis was performed on pTrc99AtaqTth(N-D) DNA using the mutagenic primer 240-60-08 5'-gcc gag gag atc cgc cgc ctc gag gcc-3' (SEQ ID NO:284)

to introduce the A468R mutation.

**9. Construction of TaqTth(N-D) A472E (DNA sequence SEQ ID NO:91;
amino acid sequence SEQ ID NO:24)**

Site specific mutagenesis was performed on pTrc99AtaqTth(N-D) DNA using the mutagenic primer 240-60-09 5'-gcc cgc ctc gag gag gag gtc ttc cgc-3' (SEQ ID NO:285)

to introduce the A472E mutation.

**10. Construction of TaqTth(N-D) G499R (DNA sequence SEQ ID NO:92;
amino acid sequence SEQ ID NO:25)**

Site specific mutagenesis was performed on pTrc99AtaqTth(N-D) DNA using the mutagenic primer 240-60-10 5'-ttt gac gag cta agg ctt ccc gcc atc-3' (SEQ ID NO:286) to

introduce the G499R mutation.

**11. Construction of TaqTth(N-D) E507Q (DNA sequence SEQ ID NO:93;
amino acid sequence SEQ ID NO:26)**

Site specific mutagenesis was performed on pTrc99AtaqTth(N-D) DNA using the mutagenic primer 276-046-04 5'-atc gcc aag acg caa aag acc ggc aag-3' (SEQ ID

NO:287) to introduce the E507Q mutation.

12. Construction of TaqTth(N-D) Y535H (DNA sequence SEQ ID NO:94; amino acid sequence SEQ ID NO:27)

Site specific mutagenesis was performed on pTrc99AtaqTth(N-D) DNA using the mutagenic primer 240-60-11 5'-aag atc ctg cag cac cgg gag ctc acc-3' (SEQ ID NO:288) to introduce the Y535H mutation.

13. Construction of TaqTth(N-D) S543N (DNA sequence SEQ ID NO:95; amino acid sequence SEQ ID NO:28)

Site specific mutagenesis was performed on pTrc99AtaqTth(N-D) DNA using the mutagenic primer 240-60-12 5'-acc aag ctg aag aac acc tac att gac-3' (SEQ ID NO:289) to introduce the S543N mutation.

14. Construction of TaqTth(N-D) I546V (DNA sequence SEQ ID NO:96; amino acid sequence SEQ ID NO:29)

Site specific mutagenesis was performed on pTrc99AtaqTth(N-D) DNA using the mutagenic primer 240-60-13 5'-aag agc acc tac gtg gac ccc ttg ccg-3' (SEQ ID NO:290) to introduce the I546V mutation.

15. Construction of TaqTth(N-D) D551S/I553V (DNA sequence SEQ ID NO:97; amino acid sequence SEQ ID NO:30)

Site specific mutagenesis was performed on pTrc99AtaqTth(N-D) DNA using the mutagenic primer 240-60-14 5'-att gac ccc ttg ccg agc ctc gtc cac ccc agg acg ggc-3' (SEQ ID NO:291) to introduce the D551S and the I553V mutations.

16. Construction of TaqDN RX HT W417L/G418K/E507Q (DNA sequence SEQ ID NO:98; amino acid sequence SEQ ID NO:31)

The TaqDN RX HT W417L/G418K/E507Q triple mutant was made by combining the TaqTth(D-B)W417L/G418K with the TaqTth(N-D) E507Q. TaqTth(D-B)W417L/G418K was cut with the restriction enzymes NdeI and BamHI, and the larger, vector fragment was isolated as detailed in Example 3. The TaqTth(N-D) E507Q construct was also cut with NdeI and BamHI and the smaller (approximately 795

base pairs) fragment was gel isolated and purified as detailed in Example 3. The NdeI-BamHI insert was ligated into the gel purified vector, as detailed in Example 3.

17. Construction of TaqDN RX HT W417L/E507Q (DNA sequence SEQ ID NO:99; amino acid sequence SEQ ID NO:32)

Starting with TaqDN RX HT W417L/G418K/E507Q described above, mutagenic primer 337-01-02: 5'-TTC GCC AAC CTG CTT GGG AGG CTT GAG GGG GAG -3' (SEQ ID NO:292) was used in a site specific mutagenesis reaction to change the K at amino acid position 418 back to the wild-type amino acid, G. Site specific mutagenesis was done using the Transformer Site Directed Mutagenesis Kit (Clonetech) according to the manufacturer's instructions, and as described in Experimental Example 4.

18. Construction of TaqDN RX HT G418K/E507Q (DNA sequence SEQ ID NO:100; amino acid sequence SEQ ID NO:33)

Starting with TaqDN RX HT W417L/G418K/E507Q described above, mutagenic primer 337-01-01: 5'-CTC TTC GCC AAC CTG TGG AAG AGG CTT GAG GGG -3' (SEQ ID NO:293) was used in a site specific mutagenesis reaction to change the L at amino acid position 417 back to the wild-type amino acid, W. Site specific mutagenesis was done using the Transformer Site Directed Mutagenesis Kit (Clonetech) according to the manufacturer's instructions, and as described in Experimental Example 4.

Expression and purification of mutant proteins was done as detailed in Example 3, and the cleavage activities of these proteins were characterized as describe in Example 1, part A. A comparison of the cleavage cycling rates of a selection of these mutant proteins on an RNA target is shown in Figure 14. As further discussed in the Description of the Invention, these data show that amino acids in the regions 417/418 and amino acid 507 are important in the conferring the TthPol-like RNA-dependent cleavage activity on the chimerical proteins comprising portions of TaqPol in combination with portions of TthPol that are not independently capable of providing enhanced RNA dependent activity (*i.e.*, the D-B and N-D portions of Tth). As described in the Description of the Invention, Taq DN RX HT variant carrying only the W417L, G418K and E507Q substitutions were created. By comparing their cleavage rates to that of Tth DN RX HT on the IL-6 RNA

substrate as described in Example 1, these mutations were determined to be sufficient to increase the Taq DN RX HT activity to the Tth DN RX HT level. Figure 15 shows that the Taq DN RX HT W417L/G418K/E507Q and Taq DN RX HT G418K/E507Q mutants have 1.4 times higher activity than Tth DN RX HT and more than 4 fold higher activity than Taq DN RX HT, whereas the Taq DN RX HT W417L/E507Q mutant has the same activity as the enzyme, which is about 3 fold higher than Taq DN RX HT. These results demonstrate that K418 and Q507 of TthPol are particularly important amino acids in providing RNA dependent 5' nuclease activity that is enhanced compared to TaqPol.

EXAMPLE 6

RNA-dependent 5' nuclease properties of the Taq DN RX HT G418K/E507Q 5' nuclease are similar to Tth DN RX HT with respect to salt and temperature optima

To determine if the G418K/E507Q mutations caused any significant changes in the properties of the Taq DN RX HT mutant in addition to the increased cleavage rate with the RNA target, the Taq DN RX HT G418K/E507Q (SEQ ID NO:33), Taq DN RX HT (SEQ ID NO:276), and Tth DN RX HT (SEQ ID NO:274) enzymes were compared in the RNA template dependent 5' nuclease assay under conditions where temperature and the concentrations of salt and divalent ions were varied. The upstream DNA and the template RNA strands of the substrate used in this study were linked into a single IrT molecule (SEQ ID NO:166) as shown in Figure 20A, and the labeled downstream probe (SEQ ID NO:167) was present in large excess. The 5' end of the target RNA strand was blocked with a biotin-streptavidin complex to prevent any non-specific degradation by the enzyme during the reaction (Lyamichev *et al.*, Science 260:778 [1993], Johnson *et al.*, Science 269:238 [1995]). The cleavage rates for Taq DN RX HT G418K/E507Q, Taq DN RX HT, and Tth DN RX HT are plotted as functions of temperature in Figure 20B. The closed circles represent enzyme Taq DN RX HT, the open circles represent enzyme Tth DN RX HT, and the Xs represent enzyme Taq DN RX HT G418K/E507Q. The difference in the activities of Tth and Taq DN RX HT enzymes with the IrT substrate is even greater than the difference found with the IL-6 RNA substrate when tested in a cleavage assay as described in Example 1. The G418K/E507Q mutations increase the

activity of the Taq enzyme more than tenfold and by 25% compared with the Tth enzyme. All three enzymes show a typical temperature profile of the invasive signal amplification reaction and have the same optimal temperature. No significant effect of G418K/E507Q mutations on DNA dependent 5' nuclease activity of Taq DN RX HT with the all-DNA substrate analogous to IrT (SEQ ID NO:168) under the same conditions was found.

The effects of KCl and MgSO₄ concentrations on the 5' nuclease activity of Taq DN RX HT G418K/E507Q, Taq DN RX HT, and Tth DN RX HT with the IrT substrate are shown in Figure 20C and D. The activities of all enzymes have similar salt dependencies with an optimal KCl concentration of 100 mM for Taq DN RX HT G418K/E507Q and Tth DN RX HT and 50 mM for Taq DN RX HT. The optimal MgSO₄ concentration for all enzymes is approximately 8 mM. The analysis of the data presented in Figure 20 suggests that the properties of Taq DN RX HT G418K/E507Q are much closer to those of Tth DN RX HT rather than Taq DN RX HT confirming the key role of the G418K/E507Q mutations in the recognition of the substrate with an RNA target.

To understand the mechanism of the reduction of the 5' nuclease activity in the presence of an RNA versus a DNA target, the Michaelis constant (K_m) and the maximal catalytic rate (k_{cat}) of all three enzymes were determined, using an excess of the IrT substrate (SEQ ID NO:166) and the downstream probe (SEQ ID NO:167) and a limiting enzyme concentration. For these measurements, ten- μ l reactions were assembled containing 10 mM MOPS, pH 7.5, 0.05% Tween 20, 0.05% Nonidet P-40, 10 μ g/ml tRNA, 4 mM MgCl₂, 1 nM of enzyme (Taq DN RX HT, Tth DN RX HT, or Taq DN RX HT G418K/E507Q) and different concentrations (0.125, 0.25, 0.5 or 1 μ M) of an equimolar mixture of the IrT target and the downstream probe. The cleavage kinetics for each enzyme and each substrate concentration were measured at 46°C. Reactions were stopped by the addition of 10 μ l of 95% formamide containing 10 mM EDTA and 0.02% methyl violet (Sigma). One μ l of each stopped reaction digest was fractionated on a 20% denaturing acrylamide gel (19:1 cross-linked), with 7M urea, and in a buffer of 45 mM Tris-borate, pH 8.3, 1.4mM EDTA. Gels were scanned on an FMBIO-100 fluorescent gel scanner (Hitachi) using a 585 nm filter. The fraction of cleaved product (determined

from intensities of bands corresponding to uncut and cut substrate with FMBIO Analysis software, version 6.0, Hitachi) was plotted as a function of reaction time. The initial cleavage rates were determined from the slopes of linear part of the cleavage kinetics and were defined as the concentration of cut product divided by the enzyme concentration and the time of the reaction (in minutes). The Michaelis constant K_m and the maximal catalytic rate k_{cat} of each enzyme with IrT substrate were determined from the plots of the initial cleavage rate as functions of the substrate concentration.

It was found that all three enzymes have similar K_m values (in the range of 200-300 nM) and k_{cat} values of approximately 4 min⁻¹ for Taq DN RX HT and Tth DN RX HT and of 9 min⁻¹ for Taq DN RX HT G418K/E507Q. That the G418K/E507Q mutations increase the k_{cat} of Taq DN RX HT more than two fold, but have little effect on K_m suggest that the mutations position the substrate in an orientation more appropriate for cleavage, rather than simply increase the binding constant.

EXAMPLE 7

Use of molecular modeling to further improve RNA-dependent 5' nuclease activity

A. Point mutants

To develop enzymes with altered function, sequence changes were introduced by site-specific mutagenesis in predetermined locations or by random mutagenesis. Locations for site-specific mutagenesis were chosen based on evidence from chimeric studies, relevant published literature, and molecular modeling. Seven additional mutant enzymes were developed from the Tth DN RX HT enzyme, and twenty additional mutant enzymes were developed from the Taq DN RX HT enzyme, both discussed previously. Some of the mutant enzymes are the result of multiple mutagenesis reactions, that is, more than one change has been introduced to obtain the final product. Mutation reactions were done using the Tth DN RX HT construct (SEQ ID NO:273) described in Example 2C2, or the Taq DN RX HT construct (SEQ ID NO:275), described in Example 2C1 unless otherwise stated. Plasmid DNA was purified from 200 ml of JM109 overnight culture using QIAGEN Plasmid Maxi Kit (QIAGEN, Chatsworth, CA) according to the manufacturer's protocol to obtain enough starting material for all mutagenesis reactions.

All site-specific mutations were introduced using the Transformer Site Directed mutagenesis Kit (Clontech) according to the manufacturer's protocol. One of two different selection primers, Trans Oligo AlwNI/SpeI or Switch Oligo SpeI/AlwNI (Clontech, Palo Alto CA catalog #6488-1 or catalog #6373-1) was used for all mutagenesis reactions described. The selection oligo used in a given reaction is dependent on the restriction site present in the vector. All mutagenic primers for both the site-specific mutagenesis and the random mutagenesis were synthesized by standard synthetic chemistry. Resultant colonies for both types of reactions were *E.coli* strain JM109. Random mutagenesis methods are described below.

Mutants were tested via the rapid screening protocol detailed in Example 1. Then, if more detailed analysis was desired, or if a larger protein preparation was required, expression and purification of mutant proteins was done as detailed in Example 3.

1. Construction of Tth DN RX HT H641A, Tth DN RX HT H748A, Tth DN RX HT H786A

Site specific mutagenesis was performed on pTrc99A Tth DN RX HT DNA using the mutagenic primer 583-001-02: 5'-gct tgc ggt ctg ggt ggc gat gtc ctt ccc ctc-3' (SEQ ID NO:294) to introduce the H641A mutation (DNA sequence SEQ ID NO:101; amino acid sequence SEQ ID NO:34), or the mutagenic primer 583-001-03: 5' cat gtt gaa ggc cat ggc ctc cgc ggc ctc cct-3' (SEQ ID NO:295) to generate the H748A mutant (DNA sequence SEQ ID NO:102; amino acid sequence SEQ ID NO:35), or the mutagenic primer 583-001-04: 5'-cag gag gag ctc gtt ggc gac ctg gag gag-3' (SEQ ID NO:296) to generate the H786A mutant enzyme (DNA sequence SEQ ID NO:103; amino acid sequence SEQ ID NO:36).

2. Construction of Tth DN RX HT (H786A/G506K/Q509K)

Starting with the mutant Tth DN RX HT H786A, generated above, site specific mutagenesis was done using the mutagenic primer 604-022-02: 5'-gga gcg ctt gcc tgt ctt ctt cgt ctt ctt caa ggc ggg agg cct-3' (SEQ ID NO:297) to generate this variant termed "TthAKK", (DNA sequence SEQ ID NO:104; amino acid sequence SEQ ID NO:37).

3. Construction of Taq DN RX HT (W417L/G418K/E507Q/H784A)

Mutagenic oligonucleotide 158-029-02: 5'-gag gac cag ctc gtt ggc gac ctg aag gag cat-3' (SEQ ID NO:298) was used in a site specific mutagenesis reaction to introduce the H784A mutation and generate this construct termed "Taq4M" (DNA sequence SEQ ID NO:105; amino acid sequence SEQ ID NO:38).

4. Construction of Taq4M H639A, Taq4M R587A, Taq4M G504K and Taq4M G80E

Site specific mutagenesis was done on the Taq4M mutant, using primer 473-010-11: 5'-gaggggcgggacatcgccacggagaccgccagc-3' (SEQ ID NO:299) to generate the Taq 4M H639A mutant (DNA sequence SEQ ID NO:106; amino acid sequence SEQ ID NO:39), primer 473-010-10: 5'-cag aac atc ccc gtc gcc acc ccg ctt ggg cag-3' (SEQ ID NO:300) to generate Taq 4M R587A (DNA sequence SEQ ID NO:107; amino acid sequence SEQ ID NO:40), primer 300-081-06: 5'-ggg ctt ccc gcc atc aag aag acg gag aag acc-3' (SEQ ID NO:301) to generate Taq 4M G504K (DNA sequence SEQ ID NO:108; amino acid sequence SEQ ID NO:41), and primer 330-088-04: 5'-cta ggg ctt ccc gcc atc aag aag acg caa aag acc ggc-3' (SEQ ID NO:302) to generate the Taq 4M G80E mutant (DNA sequence SEQ ID NO:109; amino acid sequence SEQ ID NO:42).

5. Construction of Taq 4M P88E/P90E and Taq 4M L109F/A110T

Starting with Taq 4M described above, site specific mutagenesis was done using primer 473-087-03: 5'-ccg ggg aaa gtc ctc ctc cgt ctc ggc ccg gcc cgc ctt-3' (SEQ ID NO:303) to generate the P88E/P90E mutations (DNA sequence SEQ ID NO:110; amino acid sequence SEQ ID NO:43), or primer 473-087-05: 5'-cgg gac ctc gag gcg cgt gaa ccc cag gag gtc cac-3' (SEQ ID NO:304) to generate the L109F/A110T mutations (DNA sequence SEQ ID NO:111; amino acid sequence SEQ ID NO:44).

6. Construction of Taq DN RX HT

(W417L/G418K/G499R/A502K/I503L /G504K/ E507K/H784A)

Two PCR reactions were performed, first using construct Taq4M (Taq W417L/G418K/G504K/E507Q/H784A) as a template. Using primers 158-84-01 5'-CTCCTCCACGAGTTCGGC-3' (SEQ ID NO:305) and 535-33-02 5'-ACC GGT CTT CTT CGT CTT CTT CAA CTT GGG AAG CCT GAG CTC GTC AAA-3' (SEQ ID NO:306) a 620 base pair PCR fragment was generated. Another 510 base pair PCR product was generated using primer 535-33-01 5'-AAG ACG AAG AAG ACC GGT

AAG CGC TCC ACC AGC-3' (SEQ ID NO:307) and 330-06-03 5'-GTC GAC TCT
AGA TCA GTG GTG GTG GTG GTG GTG CTT GGC CGC CCG GCG CAT C-3'
(SEQ ID NO:308). The two PCR products overlap such that a final recombinant PCR
amplification was done using the outside primers 158-84-01 and 330-06-03 to yield the
5 1182 base pair product. The recombinant PCR product was digested with the restriction
enzymes NotI and BamHI according to the manufacturer's instructions to yield a 793
base pair fragment. The parent plasmid Taq4M was also digested with the same enzymes
and used as the vector for ligation. All DNA fragments were TAE agarose gel purified
prior to ligation. The fragment was ligated into the vector, and transformed into JM109
10 cells, thus incorporating the mutations G499R, A502K, I503L, and E507K as well as the
restriction endonuclease site, AgeI. This construct is termed "Taq 8M" (DNA sequence
SEQ ID NO:112; amino acid sequence SEQ ID NO:45).

B. Random Mutagenesis

15 Numerous enzymes with altered function were generated via random
mutagenesis. The regions of the protein targeted for random mutagenesis were chosen
based on molecular modeling data and from information in the literature. Different
mutagenic primers were used to introduce mutations into different regions of the protein.
Random mutagenesis was performed on the Taq variant Taq 4M G504K (Taq DN RX
20 HT W417L/G418K/G504K/E507Q/H784A/) (SEQ ID NO:108) described above and
mutant PCR fragments generated in the mutagenesis reaction were exchanged for
homologous regions in Taq8M (SEQ ID NO:112) unless otherwise stated.

Random mutagenesis was also performed on the Tth DN RX HT H786A (SEQ ID
NO:103) described above. Mutant PCR fragments generated with the Tth DN RX HT
25 H786A template were exchanged for homologous regions in the unaltered Tth DN RX
HT H786A.

1. Random mutants in amino acid residues 500-507 or 513-520

The first mutagenic oligonucleotide, 535-054-01: 5'-gga gcg ctt acc ggt ctt (ttg
30 cgt ctt ctt gat ctt ggg aag) cct tag ctc gtc aaa gag-3' (SEQ ID NO:309) was used in
conjunction with 158-84-01: 5'-CTC CTC CAC GAG TTC GGC-3' (SEQ ID NO:310) to

install random residues from amino acid position 500 to 507 of Taq polymerase variant Taq DN RX HT W417L/G418K/G504K/E507Q/H784A (SEQ ID NO:108). This was accomplished by synthesizing the primer 535-054-01 such that only 91% of the bases within the parenthesis are unaltered while the remaining 9% of the bases are an equal mixture of the other 3 nucleotides. The initial, unaltered sequence of this oligo includes the G499R, A502K and the Q507K changes.

To generate mutations in the region 500-507, primer 535-054-01 and primer 158-84-01 were used in a PCR reaction, using the Advantage cDNA PCR kit (Clontech) and Taq variant described above, as the target. This PCR fragments was then run on a 1% TAE agarose gel, excised and purified with QIAquick Gel Extraction Kit (Qiagen, Valencia CA, catalog # 28706). The purified fragment was cut with NotI and AgeI and ligated into pTaq8M that had been linearized with NotI and AgeI. JM109 *E.coli* cells (Promega) were transformed with the ligated products. Clones were tested as described below.

The second mutagenic oligonucleotide (used in a separate reaction) 535-054-02: 5'-caa aag acc ggt aag cgc (tcc acc agc gcc gcc gtc ctg gag) gcc ctc cgc gag gcc cac-3' (SEQ ID NO:311) was used in conjunction with 330-06-03: 5'-GTC GAC TCT AGA TCA GTG GTG GTG GTG GTG GTG CTT GGC CGC CCG GCG CAT C-3' (SEQ ID NO:312) to install random residues from amino acid 513-520. The bases within the parenthesis of primer 535-054-02 are also 91% wild-type and 3% each of the other 3 nucleotides.

To generate mutations in the region 513-520, primer 535-054-02 and primer 535-054-02 were used in a PCR reaction with Taq DN RX HT W417L/G418K/G504K/E507Q/H784A (SEQ ID NO:108) as template, as described above. The resulting PCR fragment was purified as above and cut with the restriction enzymes AgeI and BamHI. The cut fragment was then ligated into the Taq8M construct, also linearized with AgeI and BamHI. JM109 *E.coli* cells were transformed with the ligated products. Clones were tested as described Example 1. Mutants developed from these include:

Taq DN RX HT W417L/G418K/G499R/A502K/K504N/E507K/H784A (M1-13) (DNA sequence SEQ ID NO:113; amino acid sequence SEQ ID NO:46).

Taq DN RX HT W417L/G418K/G499R/L500I/A502K/G504K/Q507H/H784A (M1-36)
(DNA sequence SEQ ID NO:114; amino acid sequence SEQ ID NO:47).

Taq DN RX HT W417L/G418K/G499R/A502K/I503L/G504K/E507K/T514S/H784A
(M2-24) (DNA sequence SEQ ID NO:115; amino acid sequence SEQ ID NO:48).

5 Taq DN RX HT W417L/G418K/G499R/A502K/I503L/G504K/E507K/ V518L/H784A
(M2-06) (DNA sequence SEQ ID NO:116; amino acid sequence SEQ ID NO:49).

2. TthDN RX HT H786A random mutagenesis

To generate mutants in the helix-hairpin-helix region of the TthDN RX HT
H786A (SEQ ID NO:36) enzyme, two different PCR reactions were performed using the
10 H786A (SEQ ID NO:103) mutant as a template. The two PCR products overlap such that
a recombinant PCR reaction can be performed (Higuchi, in PCR Technology, H. A.
Erlich, ed., Stockton Press, New York. pp61-70 [1989]). This final PCR product is then
exchanged with the homologous region of the TthDN H786A mutant by using restriction
enzyme sites located on the ends of the fragment and within the TthDN H786A sequence.

15 Starting with TthDN H786A discussed above, and using primer 604-08-06: 5'-gtc
gga ggg gtc ccc cac gag-3' (SEQ ID NO:313) and primer 390-76-08: 5'-tgt gga att gtg
agc gg (SEQ ID NO:314), a 620 base pair PCR fragment was generated. PCR reactions
were performed using the Advantage cDNA PCR kit (Clontech) according to
manufacturer's instructions. This PCR product includes amino acids 1-194. No
20 mutations were introduced via this reaction, however the restriction enzyme site EcoRI is
present at the 5' end.

Starting with TthDN RX HT H786A discussed above, and using mutagenic
primer 604-08-05: 5'-ctc gtg ggg gac ccc tcc gac aac ctc (ccc ggg gtc aag ggc atc ggg
gag aag acc gcc) ctc aag ctt ctc aag-3' (SEQ ID NO:315) and primer 209-74-02: 5'-gtg
25 gcc tcc ata tgg gcc agg ac-3' (SEQ ID NO:316) a 787 base pair PCR fragment was
generated. PCR reactions were done as above. This fragment does contain random
mutations, due to the presence of the mutagenic primer, 604-08-05. The bases within the
parenthesis of this primer were synthesized such that 91% of the sequence is wild-type,
while the additional 9% is evenly divided between the remaining 3 bases.

30 The two PCR fragments overlap, and were combined in a recombinant PCR
reaction. Primers 390-76-08 and 209-74-02 were added, and the Advantage cDNA PCR

kit (Clontech) was again used according to manufacturer's instructions. A 1380 base pair product was generated from this reaction.

The recombinant PCR product was cut with the restriction enzymes EcoRI and NotI according to the manufacturer's instructions to yield a 986 base pair fragment.

5 TthDN RX HT H786A was prepared by cutting with the same enzymes. The fragment was then ligated into the vector, and transformed into JM109 cells. New mutants developed from this set of reactions include:

TthDN RX HT H786A/P197R/K200R (DNA sequence SEQ ID NO:117; amino acid
10 sequence SEQ ID NO:50).

TthDN RX HT H786A/K205Y (DNA sequence SEQ ID NO:118; amino acid sequence
SEQ ID NO:51).

TthDN RX HT H786A/G203R (DNA sequence SEQ ID NO:119; amino acid sequence
SEQ ID NO:52).

15 3. Construction of Taq DN RX HT W417L/G418K/H784A

L109F/A110T/G499R/A502K/I503L/G504K/E507K/T514S (Taq SS)

Starting with Taq DN RX HT

W417L/G418K/G499R/A502K/I503L/G504K/E507K/T514S/H784A (SEQ ID NO:115)
mutant described above, primer 473-087-05: 5'-cgg gac ctc gag gcg cgt gaa ccc cag gag
20 gtc cac-3' (SEQ ID NO:317) was used in conjunction with the appropriate selection
primer in a site specific mutagenesis reaction to incorporate the L109F and A110T
mutations to generate this enzyme, termed "TaqSS" (DNA sequence SEQ ID NO:120;
amino acid sequence SEQ ID NO:53).

25 4. Construction of Taq DN RX HT W417L/G418L/H784A

P88E/P90E/G499R/A502K/I503L/G504K/E507K/T514S

Starting with Taq DN RX HT

W417L/G418K/G499R/A502K/I503L/G504K/E507K/T514S/H784A (SEQ ID NO:115)
mutant described above, primer 473-087-03: 5'-ccg ggg aaa gtc ctc ctc cgt ctc ggc ccg
gcc cgc ctt-3' (SEQ ID NO:318) was used in conjunction with the appropriate selection
30 primer in a site specific mutagenesis reaction to incorporate the P88E and P90E

mutations to generate this enzyme (DNA sequence SEQ ID NO:121; amino acid sequence SEQ ID NO:54).

5. TaqSS random mutagenesis

Random mutagenesis was used to introduce additional changes in the helix-hairpin-helix domain of the TaqSS mutant (SEQ ID NO:120). The mutagenesis was done as described in example 9 above. In the first step, two different but overlapping PCR products were generated. One of the PCR products, generated with oligos 390-76-08 (SEQ ID NO:314), and 604-08-04: 5'-gtc gga ctc gtc acc ggt cag ggc-3' (SEQ ID NO:319) incorporates the EcoRI site into the fragment, but does not incorporate any mutations. The second PCR product utilizes mutagenic primer 604-08-03: 5'-ctg acc ggt gac gag tcc gac aac ctt (ccc ggg gtc aag ggc atc ggg gag aag acg gcg) agg aag ctt ctg gag-3' (SEQ ID NO:320) and primer 209-74-02 (SEQ ID NO:316). This fragment contains random point mutations, and when combined via recombinant PCR with the first fragment, can be cut with the restriction enzymes EcoRI and NotI, and ligated into the TaqSS construct, also cut with EcoRI and NotI. The ligated construct was then transformed into JM109. Colonies were screened as described below. Enzymes developed from this mutagenesis include:

TaqSS K198N (DNA sequence SEQ ID NO:112; amino acid sequence SEQ ID NO:55).

TaqSS A205Q (DNA sequence SEQ ID NO:123; amino acid sequence SEQ ID NO:56).

TaqSS I200M/A205G (DNA sequence SEQ ID NO:124; amino acid sequence SEQ ID NO:57).

TaqSS K203N (DNA sequence SEQ ID NO:125; amino acid sequence SEQ ID NO:58).

TaqSS T204P (DNA sequence SEQ ID NO:126; amino acid sequence SEQ ID NO:59).

6. Construction of TaqSS R677A

To generate enzymes with sequence changes in both the arch region and in the polymerase region, additional specific point mutations were generated in TaqSS. Site specific mutagenesis was performed as described above using the oligo 473-060-10:

5'-tag ctc ctg gga gag ggc gtg ggc cga cat gcc-3' (SEQ ID NO:321) to generate the

TaqSS R677A mutant (DNA sequence SEQ ID NO:127; amino acid sequence SEQ ID NO:60).

7. Construction of TaqTthAKK (DNA sequence SEQ ID NO:128; amino acid sequence SEQ ID NO:61) and TthTaq5M (DNA sequence SEQ ID NO:129; amino acid sequence SEQ ID NO:62)

Chimeric mutant TaqTthAKK and TthTaq5M were generated by cutting Tth DN
RX HT (H786A/G506K/Q509K) (SEQ ID NO:104; here abbreviated TthAKK) or Taq
4M G504 (SEQ ID NO:108; here abbreviated Taq 5M) with the restriction endonucleases
EcoRI and NotI. The smaller insert fragments as well as the larger vector fragments were
gel purified as detailed in Example 3D, and the insert fragments were exchanged between
the two mutants and ligated as described in Example 3D. Screening and verification of
the construct sequence was also done as in Example 3D.

EXAMPLE 8

Improvement of RNA-dependent 5' nuclease activity in other polymerases

Information gained from the TaqPol/TthPol recombinations, mutagenesis and
modeling, was used to make comparable mutations in additional DNA polymerases and
examined the effects on the cleavage activities of these enzymes. The DNA polymerases
of *Thermus filiformus* (TfiPol) and *Thermus scotoductus* (TscPol) were cloned and
purified as described in Example 2. The mutagenesis of these two proteins is described
below.

A. Construction of TfiPolDN2M

Mutagenesis of pTrc99a-TfiPol (SEQ ID NO:249) was done using the
QuikChange site-directed mutagenesis kit (Stratagene) according to the manufacturer's
protocol. The P420K mutation was made with the following two oligonucleotides;
5'-CTTCCAGAACCTCTTTAAACGGCTTCCGAGAAG (SEQ ID NO:322) and
5'-CTTCTCGGAAAGCCGTTTAAAGAGGTTCTGGAAG (SEQ ID NO:323). The
E507Q mutation was made with the following two oligonucleotides;
5'-CCGGTGGGCCGACGCAGAAGACGGGCAAGC (SEQ ID NO:324) and
5'-GCTTGCCCGTCTTCTGCGTCCGGCCCACCGG (SEQ ID NO:325). The D785N
mutation was made with the following two oligonucleotides;
5'-CTCCTCCAAGTGCACAACGAGCTGGTCCTGG (SEQ ID NO:326) and

5'-CCAGGACCAGCTCGTTGTGCACTTGGAGGAG (SEQ ID NO:327). The plasmid containing all three mutations is called pTrc99a-TfiPolDN2M, (DNA sequence SEQ ID NO:130; amino acid sequence SEQ ID NO:63).

B. Construction of TscPolDN2M

5 Mutagenesis of pTrc99a-TscPol (SEQ ID NO:253) was done with the QuikChange site-directed mutagenesis kit (Stratagene) according to the manufacturer's protocol. The E416K mutation was made with the following two oligonucleotides; 5'-GCCGCCCTCCTGAAGCGGCTTAAGGG (SEQ ID NO:328) and 5'-CCCTTAAGCCGCTTCAGGAGGGCGGC (SEQ ID NO:329). The E505Q mutation 10 was made with the following two oligonucleotides; 5'-ATCGGCAAGACGCAGAAGACGGGCAAGC (SEQ ID NO:330) and 5'-GCTTGCCCGTCTTCTGCGTCTTGCCGAT (SEQ ID NO:331). The D783N mutation was made with the following two oligonucleotides; 5'-TTGCAGGTGCACAACGAAGTGGTCCTC (SEQ ID NO:332) and 15 5'-GAGGACCAGTTCGTTGTGCACCTGCAA (SEQ ID NO:333). The plasmid containing all three mutations is called pTrc99a-TscPolDN2M, (DNA sequence SEQ ID NO:131; amino acid sequence SEQ ID NO:64).

C. Chimerics of Tsc, Tfi, Tth and Taq mutants

1. **Construction of TfiTth AKK (DNA sequence SEQ ID NO:132; amino acid sequence SEQ ID NO:65), TscTthAKK (DNA sequence SEQ ID NO:133; amino acid sequence SEQ ID NO:66), TfiTaq5M (DNA sequence SEQ ID NO:134; amino acid sequence SEQ ID NO:67) and TscTaq5M (DNA sequence SEQ ID NO:135; amino acid sequence SEQ ID NO:68)**

25 To generate chimeric enzymes between Tth DN RX HT (H86A/G506K/Q509K) (here abbreviated TthAKK, SEQ ID NO:104) or Taq 4M G504 (here abbreviated Taq 5M, SEQ ID NO:108), and Tfi DN 2M (SEQ ID NO:130), or Tsc DN 2M (SEQ ID NO:131), additional restriction endonuclease sites were introduced by site specific mutagenesis into the named Tfi and Tsc mutants. Mutagenic primers 700-011-01 5'-cag acc atg aat tcc acc cca ctt ttt gac ctg gag-3' (SEQ ID NO:334) and 700-011-02 5'-gtg gac 30 gcg gcc gcc cga ggc cgc cgc cag ggc cag-3' (SEQ ID NO:335) were used to introduce an

EcoRI site at amino acid position 1 and a NotI site at amino acid position 331 in Tfi DN 2M. Mutagenic primers 700-011-03 5'-cag acc atg aat tcc ctg ccc ctc ttt gag ccc aag-3' (SEQ ID NO:336) and 700-011-04 5'-gta aac cgc gcc gcc cca ggc ggc ggc caa ggc gtt-3' (SEQ ID NO:337) were used to introduce an EcoRI site at amino acid position 1 and a NotI site at amino acid position 327 in Tsc DN 2M. PCR reactions were done using the Advance cDNA PCR kit (Clonotech) according to manufacturer's instructions and either Tfi DN 2M or Tsc DN 2M as target, with their corresponding primers. The 1017 base pair PCR products were cut with both EcoRI and NotI to yield 993 base pair insert fragments that were gel purified as described in Example 3D. The mutants Taq4M G504K (SEQ ID NO:108) and Tth DN RX HT (H786A/G506K/Q509K) (SEQ ID NO:104) were also cut with EcoRI and NotI, and the larger, vector fragment was gel isolated as above. Ligations were performed as detailed in Example 3D, as was the screening and verification of the new constructs.

EXAMPLE 9

Additional Enzymes having Improved RNA-dependent 5' nuclease activity

Generation of Tfi DN 2M(Δ N)

To facilitate later cloning steps, an endogenous restriction enzyme site (Not I) was removed from the polymerase region of the TfiPolDN2M gene (SEQ ID NO: 130 described in Example 8A), and a unique Not I site was inserted in a more advantageous position.

The endogeneous Not I site was removed as follows. The QUIKCHANGE Site-Directed Mutagenesis Kit from (Stratagene) was used according to manufacturer's instructions with the mutagenic primers 5'-gag-gtg-gag-cgg-ccc-ctc-tcc-cgg-gtc-ttg (SEQ ID NO: 338) and 5'-caa-gac-ccg-gga-gag-ggg-ccg-ctc-cac-ctc (SEQ ID NO: 339). The new construct was named Tfi DN 2M(Δ N) (DNA sequence SEQ ID NO: 340; amino acid sequence SEQ ID NO: 341).

Generation of Tfi DN 2M(N), Tsc DN 2M(N)

To install a unique NotI site (at amino acid position 328) in Tfi DN 2M(Δ N) (SEQ ID NO: 340), primers 886-088-07 (SEQ ID NO: 342) 5'-tgg-cgg-cgg-cct-cgg-gcg-gcc-gcg-tcc-acc-ggg-caa-ca-3' and 700-010-03 5'-ctt-ctc-tca-tcc-gcc-aaa-aca-gcc (SEQ ID NO: 343) were used in a PCR reaction with Tfi DN 2M(Δ N) (SEQ ID NO: 340) as template. The resulting PCR fragment was purified and cut with the restriction enzymes NotI and SalI. The cut fragment was then ligated into the TfiTthAKK (SEQ ID NO: 132, described in example 8,C) construct which was also digested with NotI and SalI. The new construct was termed Tfi DN 2M(N) (DNA sequence SEQ ID NO: 345; amino acid sequence SEQ ID NO: 346).

To introduce a Not I restriction endonuclease site into mutant TscPol DN 2M (previously described in Example 8B above), PCR was performed with primers: 886-088-05 (SEQ ID NO: 344) 5'-tgg-ccg-ccg-cct-ggg-gcg-gcc-gcg-ttt-acc-ggg-cgg-ag-3' and 700-010-03 (SEQ ID NO: 343) 5'-ctt-ctc-tca-tcc-gcc-aaa-aca-gcc-3' using TscPolDN2M (SEQ ID NO: 131) as template. The NotI-SalI digested fragment of the PCR product was then sub-cloned into NotI and SalI digested TscTthAKK vector (SEQ ID NO: 133). The resulting construct was termed Tsc DN 2M(N) (DNA sequence SEQ ID NO:347; amino acid sequence SEQ ID NO:348).

Generation of Tfi DN 2M(N)AKK and Tsc DN2M(N)AKK

To generate the "AKK" set of mutations (G504K/E507K/H784A/D785N), site specific mutagenesis was done on the Tfi DN 2M(N) mutant (DNA sequence SEQ ID NO: 345), using primers 959-022-01 to -04: 5'-ggc-ctc-acc-ccg-gtg-aag-cgg-acg-aag-aag-acg-ggc-aag-cgc-3', 5'-gcg-ctt-gcc-cgt-ctt-ctt-cgt-ccg-ctt-cac-cgg-ggt-gag-gcc-3', 5'-ctc-ctc-ctc-caa-gtg-gcc-aac-gag-ctg-gtc-ctg-3', 5'-cag-gac-cag-ctc-gtt-ggc-cac-ttg-gag-gag-gag-3' (SEQ ID NO: 354-357) to generate the Tfi 2M(N)AKK mutant (DNA sequence SEQ ID NO: 358; amino acid sequence SEQ ID NO: 359). This construct is termed "TfiAKK".

To install the "AKK" set of mutations (G502K/E505K/H782A/D783N), into the TscDN 2M (N) construct, (DNA sequence SEQ ID NO: 347) primers 959-022-05 to -08:

5'-ggg-ctt-ccc-gcc-ata-aag-aag-acg-aag-aag-acg-ggc-aag-cgc-3', 5'-gcg-ctt-gcc-cgt-ctt-
ctt-cgt-ctt-ctt-gat-ggc-ggg-aag-ccc-3', 5'-atg-ctt-ttg-cag-gtg-gcc-aac-gaa-ctg-gtc-ctc-3',
5'-gag-gac-cag-ttc-gtt-ggc-cac-ctg-caa-aag-cat-3' (SEQ ID NO: 360-363) were used to
generate Tsc2M(N)AKK (DNA sequence SEQ ID NO: 364; amino acid sequence SEQ
ID NO: 365). This construct is termed "Tsc AKK".

Construction of point mutants by recombinant PCR

Construction of TthAKK(P195A) and TthAKK(P195K)

To introduce mutations at amino acid position 195 (either a P195A or P195K) in
the nuclease domain of TthAKK construct, mutagenic primer 785-073-01 (P195A) 5'-
ccc-tcc-gac-aac-ctc-gcc-ggg-gtc-aag-ggc-ata-3' (SEQ ID NO: 370) or 785-073-02
(P195K) 5'-ccc-tcc-gac-aac-ctc-aag-ggg-gtc-aag-ggc-ata-3' (SEQ ID NO: 371) and
primer 209-074-02: 5'-gtg-gcc-tcc-ata-tgg-gcc-agg-ac-3' (SEQ ID NO:316) were used in
a PCR reaction to generate a 787 base pair fragment. Another PCR fragment was
obtained by using the primers: 390-076-08 5'-tgt-gga-att-gtg-agc-gg-3' (SEQ ID
NO:314) and 785-073-03 5'-gag-gtt-gtc-gga-ggg-gtc-3' (SEQ ID NO: 372) in a reaction
with the same template.

The two PCR fragments overlap and were combined in a recombinant PCR
reaction. The outside primers 390-076-08 and 209-074-02 were added, and the
Advantage cDNA PCR kit (Clontech) was used according to manufacturer's instructions.
A 1380 base pair product was generated from this reaction.

The recombinant PCR product was cut with the restriction enzymes EcoRI and
NotI to yield a 986 base pair fragment. The TthAKK construct was prepared by cutting
with the same enzymes. The fragment was then ligated into the vector, and transformed
into JM109 cells. New mutants developed from this set of reactions include:
TthAKK(P195A) (DNA sequence SEQ ID NO: 373; amino acid sequence SEQ ID NO:
374) and TthAKK(P195K) (DNA sequence SEQ ID NO: 375; amino acid sequence SEQ
ID NO: 376).

Construction of TthAKK(N417K/L418K)

The same approach was used to construct TthAKK(N417K/L418K). Two
overlapping PCR fragments were generated by mutagenic primers: 785-73-07 5'-gag-

agg-ctc-cat-cgg-aag-aag-ctt-aag-cgc-ctc-gag-3' (SEQ ID NO: 377) and 700-10-03 5'-ctt-ctc-tca-tcc-gcc-aaa-aca-gcc-3' (SEQ ID NO:343), and primers 158-084-01 5'-ctc-ctc-cac-gag-ttc-ggc-3' (SEQ ID NO:310) and 785-73-08 5'-ccg-atg-gag-cct-ctc-cga-3' (SEQ ID NO: 378). The two products were combined and amplified with outside primers 700-10-03 and 158-084-01. The recombinant PCR product was cut with the restriction enzymes NotI and BamHI and ligated into the NotI/BamHI pre-cut TthAKK construct. This mutant was termed TthAKK(N417K/L418K) (DNA sequence SEQ ID NO: 379; amino acid sequence SEQ ID NO: 380).

Construction of additional TthAKK point mutants by site-directed mutagenesis

Construction of TthAKK(P255L)

Site specific mutagenesis was performed on the TthAKK construct using the mutagenic primer 886-049-05 and 886-049-06: 5'-gtg-cgc-acc-gac-ctc-ctc-ctg-gag-gtg-gac-ctc-3' (SEQ ID NO: 381), 5'-gag-gtc-cac-ctc-cag-gag-gag-gtc-ggt-gcg-cac-3' (SEQ ID NO: 382) to generate TthAKK(P255L) (DNA sequence SEQ ID NO: 383; amino acid sequence SEQ ID NO: 384).

Construction of TthAKK(F311Y)

Site specific mutagenesis was performed on the TthAKK construct using the mutagenic primer 886-049-09 and 886-049-10: 5'-ggg-gcc-ttc-gtg-ggc-tac-gtc-ctc-tcc-cgc-ccc-3' (SEQ ID NO: P385), 5'-ggg-gcg-gga-gag-gac-gta-gcc-cac-gaa-ggc-ccc-3' (SEQ ID NO: 386) to generate TthAKK(F311Y) (DNA sequence SEQ ID NO: 387; amino acid sequence SEQ ID NO: 388).

Construction of TthAKK(N221H/R224Q)

Site specific mutagenesis was performed on the TthAKK construct using the mutagenic primer 886-049-01 and 886-049-02: 5'-gaa-aac-ctc-ctc-aag-cac-ctg-gac-cag-gta-aag-cca-gaa-aac-3' (SEQ ID NO: 389), 5'-gtt-ttc-tgg-ctt-tac-ctg-gtc-cag-gtg-ctt-gag-gag-gtt-ttc-3' (SEQ ID NO: 390) to generate TthAKK(N221H/R224Q) (DNA sequence SEQ ID NO: 391; amino acid sequence SEQ ID NO: 392).

Construction of TthAKK(R251H)

Site specific mutagenesis was performed on TthAKK construct using the mutagenic primer 886-049-03 and 886-049-04: 5'-gag-ctc-tcc-cgg-gtg-cac-acc-gac-ctc-ccc-ctg-3' (SEQ ID NO: 393), 5'-cag-ggg-gag-gtc-ggt-gtg-cac-ccg-gga-gag-ctc-3' (SEQ ID NO: 394) to generate TthAKK(R251H) (DNA sequence SEQ ID NO: 395; amino acid sequence SEQ ID NO: 396).

Construction of TthAKK(P255L/R251H)

Site specific mutagenesis was performed on the TthAKK(P255L) construct using the mutagenic primer 886-088-01 and 886-088-02: 5'-gag-ctc-tcc-cgg-gtg-cac-acc-gac-ctc-ctc-ctg-3' (SEQ ID NO: 397), 5'-cag-gag-gag-gtc-ggt-gtg-cac-ccg-gga-gag-ctc-3' (SEQ ID NO: 398) to generate TthAKK(P255L/R251H) (DNA sequence SEQ ID NO: 399; amino acid sequence SEQ ID NO: 400).

Construction of Tth AKK L429V (DNA sequence SEQ ID NO: 401; amino acid sequence SEQ ID NO: 402)

Palm region randomization mutagenesis was performed on the Tth AKK construct using the mutagenic primer 680-79-02 5'-ctc cat cgg aac ctc ctt [aag cgc ctc gag ggg gag gag aag ctc ctt tgg] ctc tac cac gag gtg-3' (SEQ ID NO: 403) and reverse primer 680-79-04 5'-AAG GAG GTT CCG ATG GAG-3' (SEQ ID NO: 404) to introduce the random mutations in the Palm region. Brackets indicate a synthesis of 91% base shown and 3% all other bases.

Construction of Tth AKK E425V (DNA sequence SEQ ID NO: 405; amino acid sequence SEQ ID NO: 406)

Palm region randomization mutagenesis was performed on the Tth AKK construct using the mutagenic primer 680-79-02 5'-ctc cat cgg aac ctc ctt [aag cgc ctc gag ggg gag gag aag ctc ctt tgg] ctc tac cac gag gtg-3' (SEQ ID NO: 403) and reverse primer 680-79-04 5'-AAG GAG GTT CCG ATG GAG-3' (SEQ ID NO: 404) to introduce the random mutations in the Palm region. Brackets indicate a synthesis of 91% base shown and 3% all other bases.

Construction of Tth AKK L422N/E425K (DNA sequence SEQ ID NO: 407; amino acid sequence SEQ ID NO: 408)

Palm region randomization mutagenesis was performed on the Tth AKK construct using the mutagenic primer 680-79-02 5'-ctc cat cgg aac ctc ctt [aag cgc ctc gag ggg gag gag aag ctc ctt tgg] ctc tac cac gag gtg-3' (SEQ ID NO: 403) and reverse primer 680-79-04 5'-AAG GAG GTT CCG ATG GAG-3' (SEQ ID NO: 404) to introduce the random mutations in the Palm region. Brackets indicate a synthesis of 91% base shown and 3% all other bases.

Construction of Tth AKK L422F/W430C (DNA sequence SEQ ID NO: 409; amino acid sequence SEQ ID NO: 410)

Palm region randomization mutagenesis was performed on Tth AKK DNA using the mutagenic primer 680-79-02 5'-ctc cat cgg aac ctc ctt [aag cgc ctc gag ggg gag gag aag ctc ctt tgg] ctc tac cac gag gtg-3' (SEQ ID NO: 403) and reverse primer 680-79-04 5'-AAG GAG GTT CCG ATG GAG-3' (SEQ ID NO: 404) to introduce the random mutations in the Palm region. Brackets indicate a synthesis of 91% base shown and 3% all other bases.

Construction of Tth AKK A504F (DNA sequence SEQ ID NO: 411; amino acid sequence SEQ ID NO: 412)

Site saturation mutagenesis was performed on Tth AKK construct using the mutagenic primer 680-80-03 5'-cag gag ctt agg ctt ccc nnn ttg aag aag acg aag aag aca-3' (SEQ ID NO: 413) and reverse primer 680-80-06 5'-cct aag ctc gtc aaa gag-3' (SEQ ID NO: 414) to introduce the random mutations in the A504 amino acid.

Construction of Tth AKK A504V (DNA sequence SEQ ID NO: 415; amino acid sequence SEQ ID NO: 416)

Site saturation mutagenesis was performed on Tth AKK construct using the mutagenic primer 680-80-03 5'-cag gag ctt agg ctt ccc nnn ttg aag aag acg aag aag aca-3' (SEQ ID NO: 413) and reverse primer 680-80-06 5'-cct aag ctc gtc aaa gag-3' (SEQ ID NO: 414) to introduce the random mutations in the A504 amino acid.

Construction of Tth AKK A504S (DNA sequence SEQ ID NO: 417; amino acid sequence SEQ ID NO: 418)

Site saturation mutagenesis was performed on Tth AKK construct using the mutagenic primer 680-80-03 5'-cag gag ctt agg ctt ccc nnn ttg aag aag acg aag aag aca-3' (SEQ ID NO: 413) and reverse primer 680-80-06 5'-cct aag ctc gtc aaa gag-3' (SEQ ID NO: 414) to introduce the random mutations in the A504 amino acid.

Construction of Tth AKK S517G (DNA sequence SEQ ID NO: 419; amino acid sequence SEQ ID NO: 420)

Site saturation mutagenesis was performed on Tth AKK construct using the mutagenic primer 680-80-07 5'-GGC AAG CGC TCC ACC NNN GCC GCG GTG CTG GAG GCC CTA CGG-3' (SEQ ID NO: 421) and reverse primer 680-80-10 5'-GGT GGA GCG CTT GCC-3' (SEQ ID NO: 422) to introduce the random mutations in the S517 amino acid.

Construction of Tth AKK A518L (DNA sequence SEQ ID NO: 423; amino acid sequence SEQ ID NO: 424)

Site saturation mutagenesis was performed on Tth AKK construct using the mutagenic primer 680-80-07 5'-GGC AAG CGC TCC ACC AGC NNN GCG GTG CTG GAG GCC CTA CGG-3' (SEQ ID NO: 425) and reverse primer 680-80-10 5'-GGT GGA GCG CTT GCC-3' (SEQ ID NO: 422) to introduce the random mutations in the A518 amino acid.

Construction of Tth AKK A518R (DNA sequence SEQ ID NO: 426; amino acid sequence SEQ ID NO: 427)

Site saturation mutagenesis was performed on Tth AKK construct using the mutagenic primer 680-80-07 5'-GGC AAG CGC TCC ACC AGC NNN GCG GTG CTG GAG GCC CTA CGG-3' (SEQ ID NO: 425) and reverse primer 680-80-10 5'-ggt gga gcg ctt gcc-3' (SEQ ID NO: 422) to introduce the random mutations in the A518 amino acid.

Construction of Taq5M L451R (DNA sequence SEQ ID NO: 428; amino acid sequence SEQ ID NO: 429)

Site directed mutagenesis was performed on the Taq 5M construct using the mutagenic primer 240-60-05 5'-acg-ggg-gtg-cgc-cgg-gac-gtg-gcc-tat 3' (SEQ ID NO: 430) to introduce the L451R mutation in the Taq 5M enzyme.

Construction of Tth AKK A504K (DNA sequence SEQ ID NO: 431; amino acid sequence SEQ ID NO: 432)

Site directed mutagenesis was performed on the Tth AKK construct using the mutagenic primer 680-69-04 5'-ctt agg ctt ccc aag ttg aag aag acg aag aag aca-3' (SEQ ID NO: 433) and reverse primer 680-69-05 5'-tgt ctt ctt cgt ctt ctt caa ctt ggg aag cct aag-3' (SEQ ID NO: 434) to introduce the A504K mutation in the Tth AKK enzyme.

Construction of Tth AKK H641A (DNA sequence SEQ ID NO: 435; amino acid sequence SEQ ID NO: 436)

Site directed mutagenesis was performed on Tth AKK construct using the mutagenic primer 680-69-08 5'-gag ggg aag gac atc gcc acc cag acc gca agc-3' (SEQ ID NO: 437) and reverse primer 583-01-02 5'-get tgc ggt ctg ggt ggc gat gtc ctt ccc ctc- 3' (SEQ ID NO: 438) to introduce the H641A mutation in the Tth AKK enzyme.

Construction of Tth AKK T508P (DNA sequence SEQ ID NO: 439; amino acid sequence SEQ ID NO: 440)

Site directed mutagenesis was performed on TthAKK construct using the mutagenic primer 680-70-01 5'-ccc-gcc ttg aag aag ccg aag aag aca ggc aag-3' (SEQ ID NO: 441) and reverse primer 680-70-02 5'-ctt gcc tgt ctt ctt cgg ctt ctt caa ggc ggg-3' (SEQ ID NO: 442) to introduce the T508P mutation in the Tth AKK enzyme.

Chimeras and mutations in chimeras.

Fusion between TthAKK enzyme and alpha-peptide

A TthAKK-lacZ-alpha-peptide chimeric fusion was constructed to allow detection of mutations (including frame-shifts, deletions, insertions, etc.) which cause the inability of expression of the full-length fusion protein based on the colony blue-white screening (Wu et al., Nucleic Acids Research, 24:1710 [1996]).

Site specific mutagenesis was performed on TthAKK DNA using the mutagenic primers 959-041-03, 5'-cac-cac-cac-cac-cac-cac-gtc-gac-tag-tgc-tag-cgt-cga-cta-gct-gca-ggc-atg-caa-gct-tgg-c-3' (SEQ ID NO: 477) and 959-041-04, 5'-gcc-aag-ctt-gca-tgc-ctg-cag-cta-gtc-gac-gct-agc-act-agt-cga-cgt-ggt-ggt-ggt-ggt-g-3' (SEQ ID NO: 478) to generate pTthAKK-L with Sall site following the 6xHis tag at the C-terminus of TthAKK for the insertion of lacZ alpha peptide. The alpha peptide (of 201 amino acids) was first PCR amplified from the pCRII-TOPO vector (Invitrogen) with primers 959-041-01, 5'-cag-gaa-gcg-gcc-gcg-tcg-aca-tga-cca-tga-tta-cgc-caa-gc-3' (SEQ ID NO: 479) and 959-093-01, 5'-ggg-ccc-gcc-agg-gtc-gac-tca-ggg-cga-tgg-ccc-act-acg-tga-3' (SEQ ID NO: 480). The PCR product was then digested with restriction enzyme Sall and ligated into the pTthAKK-L vector, which was also cut with the same enzyme to generate the chimeric construct TthAKK-alpha peptide (DNA sequence SEQ ID NO: 481; amino acid sequence SEQ ID NO: 482). The orientation of the insert was confirmed by sequencing.

Construction of TthTscAKK and TthTfiAKK enzymes

The TthAKK construct was cut with the enzymes EcoRI and NotI and the smaller insert fragment was gel isolated. The TscAKK or TfiAKK constructs were also cut with EcoRI and NotI and the larger fragment was gel isolated and purified. The Tth insert (nuclease domain) was ligated into the TscAKK and TfiAKK vectors (polymerase domain) to generate TthTscAKK (DNA sequence SEQ ID NO: 447; amino acid sequence SEQ ID NO: 448) and TthTfiAKK (DNA sequence SEQ ID NO: 449; amino acid sequence SEQ ID NO: 450) chimeric constructs.

FT mutations of Tth polymerase to improve substrate specificity

Construction of Taq(FT)TthAKK

Site specific mutagenesis was performed on the TaqTthAKK construct using the mutagenic primer 473-087-05: 5'-cgg-gac-ctc-gag-gcg-cgt-gaa-ccc-cag-gag-gtc-cac-3'

(SEQ ID NO: 483) to introduce the L107F/A108T mutations and generate Taq(FT)TthAKK (DNA sequence SEQ ID NO: 484; amino acid sequence SEQ ID NO: 485).

Construction of Tfi(FT)TthAKK

Site specific mutagenesis was performed on the TfiTthAKK construct using the mutagenic primers 785-096-01: 5'-gtg-gac-ctt-ctg-ggc-ttt-acc-cgc-ctc-gag-gcc-ccg-3' (SEQ ID NO: 486) and 785-096-02: 5'-cgg-ggc-ctc-gag-gcg-ggt-aaa-gcc-cag-aag-gtc-cac-3' (SEQ ID NO: 487) to introduce the L107F/V108T mutations and generate Tfi(FT)TthAKK (DNA sequence SEQ ID NO: 349; amino acid sequence SEQ ID NO: 488).

Tfi(FT) DN 2M(N) and Tsc(FT) DN 2M(N) mutants

The L107F/V108T mutations were introduced by isolating the NotI and SalI fragment of the Tfi DN 2M(N) mutant (DNA sequence SEQ ID NO: 345) and inserting it into a NotI-SalI pre-digested Tfi(FT)TthAKK (DNA sequence SEQ ID NO: 349; amino acid sequence SEQ ID NO: 488) to yield the Tfi(FT) DN 2M(N) mutant (DNA sequence SEQ ID NO: 350, amino acid sequence SEQ ID NO: 351). To add the L107F or the E108T mutations into this Tsc-based construct, an identical procedure was done. The NotI and SalI cut fragment of Tsc DN 2M(N) mutant (DNA sequence SEQ ID NO: 348) was inserted into NotI-SalI pre-digested Tsc(FT)TthAKK (DNA sequence SEQ ID NO: 491) vector to yield Tsc(FT) DN 2M(N) mutant (DNA sequence SEQ ID NO: 352, amino acid sequence SEQ ID NO: 353).

Construction of Tfi(FT) AKK and Tsc(FT) AKK

Starting with the Tfi(FT)DN2M(N) construct described previously (DNA sequence SEQ ID NO: 350), primers (959-022-01 to -04: 5'-ggc-ctc-acc-ccg-gtg-aag-cgg-acg-aag-aag-acg-ggc-aag-cgc-3', 5'-gcg-ctt-gcc-cgt-ctt-ctt-cgt-ccg-ctt-cac-cgg-ggt-gag-gcc-3', 5'-ctc-ctc-ctc-caa-gtg-gcc-aac-gag-ctg-gtc-ctg-3', 5'-cag-gac-cag-ctc-gtt-ggc-cac-ttg-gag-gag-gag-3' (SEQ ID NO: 354-357) were used to introduce the "AKK" set of mutations (H784A, G504K and E507K) by site specific mutagenesis. The resulting

mutant construct is termed Tfi(FT)AKK (DNA sequence SEQ ID NO:366, amino acid sequence SEQ ID NO:367).

Likewise, primers 959-022-05 to -08: 5'-ggg-ctt-ccc-gcc-atc-aag-aag-acg-aag-aag-acg-ggc-aag-cgc-3', 5'-gcg-ctt-gcc-cgt-ctt-ctt-cgt-ctt-ctt-gat-ggc-ggg-aag-ccc-3', 5'-atg-ctt-ttg-cag-gtg-gcc-aac-gaa-ctg-gtc-ctc-3', 5'-gag-gac-cag-ttc-gtt-ggc-cac-ctg-caa-aag-cat-3' (SEQ ID NO: 360-363) were used in a site specific mutagenesis reaction, with the Tsc(FT)DN2M(N) mutant (DNA sequence SEQ ID NO: 352) as template to generate the Tsc(FT)AKK mutant (DNA sequence SEQ ID NO: 368; amino acid sequence SEQ ID NO: 369).

Construction of Tsc(FT)TthAKK

Site specific mutagenesis was performed on the TscTthAKK construct using the mutagenic primers 785-008-03 5'-ttt-acc-cgc-ctc-gag-gtg-ccg-ggc-3' (SEQ ID NO: 489) and reverse primer 680-21-03 5'-cgg cac ctc gag gcg ggt aaa gcc caa aag gtc cac-3' (SEQ ID NO: 490) to introduce the L107F/E108T mutations and generate Tsc(FT)TthAKK (DNA sequence SEQ ID NO: 454; amino acid sequence SEQ ID NO: 491).

Construction of Taq(FT)TscAKK and Taq(FT)TfiAKK enzymes

The Taq(FT)TthAKK construct was cut with the enzymes EcoRI and NotI and the smaller insert fragment was gel isolated. The TscAKK or TfiAKK constructs were also cut with EcoRI and NotI and the larger fragment was gel isolated and purified. The Taq(FT) insert (nuclease domain) was ligated into the TscAKK and TfiAKK vectors (polymerase domain) to generate the Taq(FT)TscAKK (DNA sequence SEQ ID NO: 443; amino acid sequence SEQ ID NO: 444) and Taq(FT)TfiAKK (DNA sequence SEQ ID NO: 445; amino acid sequence SEQ ID NO: 446) chimeric constructs.

Construction of TaqEFT-Tth(AKK) (DNA sequence SEQ ID NO:501; amino acidsequence SEQ ID NO: 502)

Site specific mutagenesis was performed on Taq(FT)-Tth(AKK) DNA (SEQ ID NO: 484) using the mutagenic primers 436-013-08: 5'-atc-gtg-gtc-ttt-gac-gcc-gag-gcc-ccc-tcc-ttc-c-3' (SEQ ID NO:503) and 436-013-09 5'-gga-agg-agg-ggg-cct-cgg-cgt-caa-aga-cca-cga-t-3' (SEQ ID NO:504) to introduce the K70E mutation.

Additional mutations to improve enzyme activity of Taq(EFT)TthAKK

Construction of TaqEFT-Tth(AKK)-A/M1 (DNA sequence SEQ ID NO:505; amino acid sequence SEQ ID NO: 506)

Site specific mutagenesis was performed on Taq(EFT)-Tth(AKK) DNA (SEQ ID NO: 501) using the mutagenic primers 1044-038-01: 5'-cag-acc-atg-aat-tcg-gag-gcg-atg-ctg-ccc-ctc-ttt-3' (SEQ ID NO:507) and 1044-038-02: (SEQ ID NO: 508) 5'-aaa-gag-ggg-cag-cat-cgc-ctc-cga-att-cat-ggt-ctg-3' to introduce the G4EA mutation.

Construction of TaqEFT-Tth(AKK)-B/M2 (DNA sequence SEQ ID NO:509; amino acid sequence SEQ ID NO: 510)

Site specific mutagenesis was performed on Taq(EFT)-Tth(AKK) DNA (SEQ ID NO: 501) using the mutagenic primers 1044-038-03: 5'-gcc-tac-cgc-acc-ttc-ttt-gcc-ctg-aag-ggc-ctc-3 and 1044-038-04: 5'-gag-gcc-ctt-cag-ggc-aaa-gaa-ggt-gcg-gta-ggc-3' (SEQ ID NO: 511 and 512, respectively) to introduce the H29F mutation.

Construction of TaqEFT-Tth(AKK)-C/M3 (DNA sequence SEQ ID NO:513; amino acid sequence SEQ ID NO: 514)

Site specific mutagenesis was performed on Taq(EFT)-Tth(AKK) DNA (SEQ ID NO: 501) using the mutagenic primers 1044-038-05: 5'-ctc-ctc-aag-gcc-ctc-aga-gag-gac-ggg-gac-gcg-3' and 1044-038-06: 5'-cgc-gtc-ccc-gtc-ctc-tct-gag-ggc-ctt-gag-gag-3' (SEQ ID NO: 515 and 516, respectively) to introduce the K57R mutation.

Construction of TaqEFT-Tth(AKK)-D/M5 (DNA sequence SEQ ID NO:517; amino acid sequence SEQ ID NO:518)

Site specific mutagenesis was performed on Taq(EFT)-Tth(AKK) DNA (SEQ ID NO: 501) using the mutagenic primers 1044-038-09: 5'-gac-gac-gtc-ctg-gcc-acc-ctg-gcc-aag-aag-gcg-3' and 1044-038-10: 5'-cgc-ctt-ctt-ggc-cag-ggt-ggc-cag-gac-gtc-gtc-3' (SEQ ID NO: 519 and 520, respectively) to introduce the S125T mutation.

Construction of TaqEFT-Tth(AKK)-E/M6 (DNA sequence SEQ ID NO:521; amino acid sequence SEQ ID NO: 522)

Site specific mutagenesis was performed on Taq(EFT)-Tth(AKK) DNA (SEQ ID NO: 501) using the mutagenic primers 1044-038-11: 5'-ggg-gag-aag-acg-gcg-ctc-aag-ctt-

ctg-gag-gag-3' and 1044-038-12: 5'-ctc-ctc-cag-aag-ctt-gag-cgc-cgt-ctt-ctc-ccc-3' (SEQ ID NO: 523 and 524, respectively) to introduce the R206L mutation.

Construction of TaqEFT-Tth(AKK)-F/M7 (DNA sequence SEQ ID NO:525; amino acid sequence SEQ ID NO: 526)

Site specific mutagenesis was performed on Taq(EFT)-Tth(AKK) DNA (SEQ ID NO: 501) using the mutagenic primers 1044-038-13: 5'-gag-ccc-gac-cgg-gag-ggg-ctt-aag-gcc-ttt-ctg-gag-agg-3' and 1044-038-14: 5'-cct-ctc-cag-aaa-ggc-ctt-aag-ccc-ctc-ccg-gtc-ggg-ctc-3 (SEQ ID NO: 527 and 528, respectively) to introduce the R269G and R271K mutations.

Construction of TaqEFT-Tth(AKK)-G/M8 (DNA sequence SEQ ID NO:529; amino acid sequence SEQ ID NO: 530)

Site specific mutagenesis was performed on Taq(EFT)-Tth(AKK) DNA (SEQ ID NO: 501) using the mutagenic primers 1044-038-15: 5'-cac-gag-ttc-ggc-ctt-ctg-gga-ggg-gag-aag-ccc-cgg-gag-gag-gcc-ccc-tgg-ccc-3' and 1044-038-16: 5'-ggg-cca-ggg-ggc-ctc-ctc-ccg-ggg-ctt-ctc-ccc-tcc-cag-aag-gcc-gaa-ctc-gtg-3' (SEQ ID NO: 531 and 532, respectively) to introduce the E290G, S291G, P292E, A294P, and L295R mutations.

Construction of TaqEFT-Tth(AKK)-H/M9 (DNA sequence SEQ ID NO:533; amino acid sequence SEQ ID NO: 534)

Site specific mutagenesis was performed on Taq(EFT)-Tth(AKK) DNA (SEQ ID NO: 501) using the mutagenic primers 1044-038-17: 5'-ctg-gcc-ctg-gcc-gcc- tgc-agg-ggc-ggc-cgc-gtg-3' and 1044-038-18: 5'-cac-gcg-gcc-gcc-cct-gca-ggc-ggc-cag-ggc-cag-3' (SEQ ID NO: 535 and 536, respectively) to introduce the A328C mutation.

Construction of TaqEFT-Tth(AKK)-I/M10 (DNA sequence SEQ ID NO:537; amino acid sequence SEQ ID NO: 538)

Site specific mutagenesis was performed on Taq(EFT)-Tth(AKK) DNA (SEQ ID NO: 501) using the mutagenic primers 1080-015-01 (SEQ ID NO:539) 5'-ggg gag aag acg gcg agg aag ctt ctg aag gag tgg ggg agc-3' and 1080-015-02 (SEQ ID NO: 540) 5'-gct ccc cca ctc ctt cag aag ctt cct cgc cgt ctt ctc ccc-3' to introduce the E210K mutation.

Construction of TaqEFT-Tth(AKK)-M1-9 (DNA sequence SEQ ID NO:541; amino acid sequence SEQ ID NO: 542)

Seven independent PCR reactions were performed, using construct TaqEFT-

Tth(AKK) (SEQ ID: 501) as a template, with the following pairs of mutagenic primers:

PCR Reaction 1; 1044-038-01 5'-cag-acc-atg-aat-tcg-gag-gcg-atg-ctg-ccc-ctc-ttt-3' (SEQ ID NO:507) and 1044-038-04 5'-gag-gcc-ctt-cag-ggc-aaa-gaa-ggt-gcg-gta-ggc-3' (SEQ ID NO:512) to yield a 108 base pair fragment, PCR Reaction 2; 1044-038-03 5'-gcc-tac-

cgc-acc-ttc-ttt-gcc-ctg-aag-ggc-ctc-3' (SEQ ID NO: 511) and 1044-038-06 5'-cgc-gtc-

ccc-gtc-ctc-tct-gag-ggc-ctt-gag-gag-3' (SEQ ID NO: 516) to yield a 117 base pair

fragment, PCR Reaction 3; 1044-038-05 5'-ctc-ctc-aag-gcc-ctc-aga-gag-gac-ggg-gac-gcg-3' (SEQ ID NO: 515) and 1044-038-10 5'-cgc-ctt-ctt-ggc-cag-ggt-ggc-cag-gac-gtc-

gtc-3' (SEQ ID NO: 520) to yield a 237 base pair fragment, PCR Reaction 4; 1044-038-09 5'-gac-gac-gtc-ctg-gcc-acc-ctg-gcc-aag-aag-gcg-3' (SEQ ID NO: 519) and 1044-038-

12 5'-ctc-ctc-cag-aag-ctt-gag-cgc-cgt-ctt-ctc-ccc-3' (SEQ ID NO: 524) to yield a 276 base pair fragment, PCR Reaction 5; 1044-038-11 5'-ggg-gag-aag-acg-gcg-ctc-aag-ctt-ctg-

gag-gag-3' (SEQ ID NO: 523) and 1044-038-14 5'-cct-ctc-cag-aaa-ggc-ctt-aag-ccc-ctc-ccg-gtc-ggg-ctc-3' (SEQ ID NO: 528) to yield a 228 base pair fragment, PCR Reaction 6;

1044-038-13 5'-gag-ccc-gac-cgg-gag-ggg-ctt-aag-gcc-ttt-ctg-gag-agg-3' (SEQ ID NO:

527) and 1044-038-16 5'-ggg-cca-ggg-ggc-ctc-ctc-ccg-ggg-ctt-ctc-ccc-tcc-cag-aag-gcc-gaa-ctc-gtg-3' (SEQ ID NO: 532) to yield a 113 base pair fragment, PCR Reaction 7;

1044-038-15 5'-cac-gag-ttc-ggc-ctt-ctg-gga-ggg-gag-aag-ccc-cgg-gag-gag-gcc-ccc-tgg-ccc-3' (SEQ ID NO: 531) and 1044-038-18 5'-cac-gcg-gcc-gcc-cct-gca-ggc-ggc-cag-ggc-

cag-3' (SEQ ID NO: 536) to yield a 157 base pair fragment. The seven PCR products

overlap such that PCR amplification in the absence of primers yielded the appropriate

1005 base pair product to introduce the K70E, G4EA, H29F, K57R, S125T, R206L,

R269G, R271K, E290G, S291G, P292E, A294P, L295R, and A328C mutations. The

product was further amplified using the outside primers: 1044-038-1 (SEQ ID NO: 507)

and 1044-038-18 (SEQ ID NO: 536) and cloned into pPCR-Script-Amp. From this

construct, the nuclease domain was again amplified using primers: 1044-038-1 (SEQ ID

NO: 507) and 1044-038-18 (SEQ ID NO: 536) and digested with EcoRI and NotI. The

digested PCR product was ligated into an EcoRI and NotI digested TaqEFT-Tth(AKK) construct and transformed into JM109 for protein expression and screening.

Construction of TaqEFT-Tth(AKK)-M1-10 (DNA sequence SEQ ID NO:543; amino acid sequence SEQ ID NO: 544)

Seven independent PCR reactions were performed, using construct TaqEFT-Tth(AKK) (SEQ ID: 501) as a template, with the following pairs of mutagenic primers: PCR Reaction 1; 1044-038-01 5'-cag-acc-atg-aat-tcg-gag-gcg-atg-ctg-ccc-ctc-ttt-3' (SEQ ID NO:507) and 1044-038-04 5'-gag-gcc-ctt-cag-ggc-aaa-gaa-ggt-gcg-gta-ggc-3' (SEQ ID NO:512) to yield a 108 base pair fragment, PCR Reaction 2; 1044-038-03 5'-gcc-tac-cgc-acc-ttc-ttt-gcc-ctg-aag-ggc-ctc-3' (SEQ ID NO: 511) and 1044-038-06 5'-cgc-gtc-ccc-gtc-ctc-tct-gag-ggc-ctt-gag-gag-3' (SEQ ID NO: 516) to yield a 117 base pair fragment, PCR Reaction 3; 1044-038-05 5'-ctc-ctc-aag-gcc-ctc-aga-gag-gac-ggg-gac-gcg-3' (SEQ ID NO: 515) and 1044-038-10 5'-cgc-ctt-ctt-ggc-cag-ggt-ggc-cag-gac-gtc-gtc-3' (SEQ ID NO: 520) to yield a 237 base pair fragment, PCR Reaction 4; 1044-038-09 5'-gac-gac-gtc-ctg-gcc-acc-ctg-gcc-aag-aag-gcg-3' (SEQ ID NO: 519) and 1080-42-02 5'-gc ttc cag gct ccc cca ctc ctt cag aag ctt gag cgc cgt ctt ctc ccc-3' (SEQ ID NO: 546) to yield a 299 base pair fragment, PCR Reaction 5; 1080-42-01 5'-ggg gag aag acg gcg ctc agg ctt ctg aag gag tgg ggg agc ctg gaa gc-3' (SEQ ID NO: 545) and 1044-038-14 5'-cct-ctc-cag-aaa-ggc-ctt-aag-ccc-ctc-ccg-gtc-ggg-ctc-3' (SEQ ID NO: 528) to yield a 228 base pair fragment, PCR Reaction 6; 1044-038-13 5'-gag-ccc-gac-cgg-gag-ggg-ctt-aag-gcc-ttt-ctg-gag-agg-3' (SEQ ID NO: 527) and 1044-038-16 5'-ggg-cca-ggg-ggc-ctc-ctc-ccg-ggg-ctt-ctc-ccc-tcc-cag-aag-gcc-gaa-ctc-gtg-3' (SEQ ID NO: 532) to yield a 113 base pair fragment, PCR Reaction 7; 1044-038-15 5'-cac-gag-ttc-ggc-ctt-ctg-gga-ggg-gag-aag-ccc-cgg-gag-gag-gcc-ccc-tgg-ccc-3' (SEQ ID NO: 531) and 1044-038-18 5'-cac-gcg-gcc-gcc-cct-gca-ggc-ggc-cag-ggc-cag-3' (SEQ ID NO: 536) to yield a 157 base pair fragment. The seven PCR products overlap such that PCR amplification in the absence of primers yielded the appropriate 1005 base pair product to introduce the K70E, G4EA, H29F, K57R, S125T, R206L, E210K, R269G, R271K, E290G, S291G, P292E, A294P, L295R, and A328C mutations. The product was further amplified using the outside primers: 1044-038-1 (SEQ ID NO: 507) and 1044-038-18 (SEQ ID NO: 536) and cloned into

pPCR-Script-Amp. From this construct, the nuclease domain was again amplified using primers: 1044-038-1 (SEQ ID NO: 507) and 1044-038-18 (SEQ ID NO: 536) and digested with EcoRI and NotI. The digested PCR product was ligated into an EcoRI and NotI digested TaqEFT-Tth(AKK) construct and transformed into JM109 for enzyme expression and screening.

K69E Mutation of FEN enzymes to further improve substrate specificity

Construction of Tth(K69E)AKK, Taq(K69E)TthAKK, Tfi(K69E)TthAKK and Tsc(K69E)TthAKK and mutants

Site specific mutagenesis was performed on TthAKK, TaqTthAKK, TfiTthAKK and TscTthAKK DNAs using the mutagenic primers 5'-atc-gtg-gtc-ttt-gac-gcc-gag-gcc-ccc-tcc-ttc-c-3' (SEQ ID NO: 492) and 5'-gga-agg-agg-ggg-cct-cgg-cgt-caa-aga-cca-cga-t-3' (SEQ ID NO: 493) to introduce the K69E mutation and to generate Tth(K69E)AKK (DNA sequence SEQ ID NO: 452; amino acid sequence SEQ ID NO: 494), Taq(K69E)ThAKK, (DNA sequence SEQ ID NO: 495; amino acid sequence SEQ ID NO: 496), Tfi(K69E)TthAKK (DNA sequence SEQ ID NO: 497; amino acid sequence SEQ ID NO: 498) and Tsc(K69E)TthAKK (DNA sequence SEQ ID NO: 499; amino acid sequence SEQ ID NO: 500) mutant enzymes.

Construction of Tsc(167-334)TthAKK

Two overlapping PCR fragments were generated by primers 390-76-08: 5'-tgt-gga-att-gtg-agc-gg-3' (SEQ ID NO:314) and 1044-041-01: 5'-ctt-ctc-tca-tcc-gcc-aaa-aca-gcc-3' (SEQ ID NO: 451) with template Tth(K69E)AKK, (DNA sequence SEQ ID NO: 452), and primers 1044-041-02: 5'-ctc-ctc-cac-gag-ttc-ggc-3' (SEQ ID NO: 453) and 209-074-02: 5'-gtg-gcc-tcc-ata-tgg-gcc-agg-ac-3' (SEQ ID NO:316) with template Tsc(K69E)TthAKK, (DNA sequence SEQ ID NO: 454). Two products were combined and amplified with outside primers 390-76-08 and 209-074-02. The recombinant PCR product was cut with the restriction enzymes EcoRI and NotI and ligated into the vector TthAKK which was prepared by cutting with the same enzymes to yield Tsc(167-

333)TthAKK construct (DNA sequence SEQ ID NO: 455; amino acid sequence SEQ ID NO: 456).

Construction of Tsc(222-334)TthAKK

Two overlapping PCR fragments were generated by primers 390-76-08: 5'-tgt-gga-att-gtg-agc-gg-3' (SEQ ID NO:314) and 1044-041-03: 5'-ttc-cag-gtg-ctt-gag-gag-gtt-ttc-cag-3' (SEQ ID NO: 457) with template Tth(K69E)AKK (DNA sequence SEQ ID NO: 452), and primers 1044-041-04: 5'-ctc-ctc-aag-cac-ctg-gaa-cag-gtg-aaa-3' (SEQ ID NO: 458) and 209-074-02: 5'-gtg-gcc-tcc-ata-tgg-gcc-agg-ac-3' (SEQ ID NO:316) with template Tsc(K69E)TthAKK, (DNA sequence SEQ ID NO: 454). Two products were combined and amplified with outside primers 390-76-08 and 209-074-02. The recombinant PCR product was cut with the restriction enzymes EcoRI and NotI and ligated into the vector TthAKK which was prepared by cutting with the same enzymes to yield Tsc(222-334)TthAKK construct (DNA sequence SEQ ID NO: 459; amino acid sequence SEQ ID NO: 460).

Construction of Tfi(222-334)TthAKK

Two overlapping PCR fragments were generated by primers 390-76-08: 5'-tgt-gga-att-gtg-agc-gg-3' (SEQ ID NO:314) and 1044-058-05: 5'-gtc-cag-gtt-ctt-gag-gag-gtt-ttc-cag-3' (SEQ ID NO: 461) with template Tth(K69E)AKK (DNA sequence SEQ ID NO: 452), and primers 1044-058-06: 5'-ctc-ctc-aag-aac-ctg-gac-cgg-gta-aag-3' (SEQ ID NO: 462) and 209-074-02: 5'-gtg-gcc-tcc-ata-tgg-gcc-agg-ac-3' (SEQ ID NO:316) with template Tfi(K69E)TthAKK (DNA sequence SEQ ID NO: 497). Two products were combined and amplified with outside primers 390-76-08 and 209-074-02. The recombinant PCR product was cut with the restriction enzymes EcoRI and NotI and ligated into the vector TthAKK which was prepared by cutting with the same enzymes to yield Tfi(222-334)TthAKK construct (DNA sequence SEQ ID NO: 463; amino acid sequence SEQ ID NO: 464).

Construction of Tfi(167-334)TthAKK

Two overlapping PCR fragments were generated by primers 390-76-08 5'-tgt-gga-att-gtg-agc-gg-3' (SEQ ID NO:314) and 1044-058-07 5'-gac-gtc-ctt-cgg-ggt-gat-gag-gtg-gcc-3' (SEQ ID NO: 465) with template Tth(K69E)AKK, (DNA sequence SEQ ID NO: 452), and primers 1044-058-08 5'-atc-acc-ccg-aag-gac-gtc-cag-gag-aag-3' (SEQ ID NO: 466) and 209-074-02 5'-gtg-gcc-tcc-ata-tgg-gcc-agg-ac-3' (SEQ ID NO:316) with template Tfi(K69E)TthAKK, (DNA sequence SEQ ID NO: 498). Two products were combined and amplified with outside primers 390-76-08 and 209-074-02. The recombinant PCR product was cut with the restriction enzymes EcoRI and NotI and ligated into the vector TthAKK which was prepared by cutting with the same enzymes to yield Tfi(167-333)TthAKK construct (DNA sequence SEQ ID NO: 467; amino acid sequence SEQ ID NO: 468).

Construction of Tsc(111-334)TthAKK

Two overlapping PCR fragments were generated by primers 390-76-08 5'-tgt-gga-att-gtg-agc-gg-3' (SEQ ID NO:314) and 1044-058-01 5'-ctc-gag-gcg-ggt-aaa-ccc-cag-gag-gtc-3' (SEQ ID NO: 469) with template Tth(K69E)AKK (DNA sequence SEQ ID NO: 452), and primers 1044-058-02 5'-ggg-ttt-acc-cgc-ctc-gag-gtg-ccc-ggc-3' (SEQ ID NO: 470) and 209-074-02 5'-gtg-gcc-tcc-ata-tgg-gcc-agg-ac-3' (SEQ ID NO:316) with template Tsc(K69E)TthAKK (DNA sequence SEQ ID NO: 499). Two products were combined and amplified with outside primers 390-76-08 and 209-074-02. The recombinant PCR product was cut with the restriction enzymes EcoRI and NotI and ligated into the vector TthAKK which was prepared by cutting with the same enzymes to yield Tsc(167-333)TthAKK construct (DNA sequence SEQ ID NO: 471; amino acid sequence SEQ ID NO: 472).

Construction of Tsc(1-167)TthAKK

Two overlapping PCR fragments were generated by primers 390-76-08 5'-tgt-gga-att-gtg-agc-gg-3' (SEQ ID NO:314) and 1044-058-03 5'-aag-cca-ctc-cgg-ggt-gat-cag-gta-acc-3' (SEQ ID NO: 473) with template Tsc(K69E)TthAKK (DNA sequence SEQ ID NO: 499), and primers 1044-058-04 5'-atc-acc-ccg-gag-tgg-ctt-tgg-gag-aag-3'

(SEQ ID NO: 474) and 209-074-02 5'-gtg-gcc-tcc-ata-tgg-gcc-agg-ac-3' (SEQ ID NO:316) with template Tth(K69E)AKK (DNA sequence SEQ ID NO: 452). Two products were combined and amplified with outside primers 390-76-08 and 209-074-02. The recombinant PCR product was cut with the restriction enzymes EcoRI and NotI and ligated into the vector TthAKK which was prepared by cutting with the same enzymes to yield Tsc(1-167)TthAKK construct (DNA sequence SEQ ID NO: 475; amino acid sequence SEQ ID NO: 476).

Modification of AfuFEN enzymes

Construction of pAfuFEN

Plasmid pAfuFEN1 was prepared as described [U.S. Patent application Ser. No. 09/684,938, WO 98/23774, each incorporated herein by reference in their entireties]. Briefly, genomic DNA was prepared from one vial (approximately 5 ml of culture) of live *A. fulgidus* bacteria from DSMZ (DSMZ #4304) with the DNA XTRAX kit (Gull Laboratories, Salt Lake City, UT) according to the manufacturer's protocol. The final DNA pellet was resuspended in 100 µl of TE (10 mM Tris HCl, pH 8.0, 1 mM EDTA). One microliter of the DNA solution was employed in a PCR using the ADVANTAGE cDNA PCR kit (Clonetech); the PCR was conducted according to manufacturer's recommendations.

The 5' end primer is complementary to the 5' end of the Afu FEN-1 gene except it has a 1 base pair substitution to create an *Nco* I site. The 3' end primer is complementary to the 3' end of the Afu FEN-1 gene downstream from the FEN-1 ORF except it contains a 2 base substitution to create a *Sal* I site. The sequences of the 5' and 3' end primers are 5'-CCGTCAACATTTACCATGGGTGCGGA-3' (SEQ ID NO:617) and 5'-CCGCCACCTCGTAGTCGACATCCTTTTCGTG (SEQ ID NO:618), respectively. Cloning of the resulting fragment was as described for the PfuFEN1 gene, U.S. Patent application Ser. No. 5,994,069, incorporated herein in its entirety for all purposes, to create the plasmid pTrc99-AFFEN1. For expression, the pTrcAfuHis plasmid was constructed by modifying pTrc99-AFFEN1, by adding a histidine tail to facilitate purification. To add this histidine tail, standard PCR primer-directed mutagenesis

methods were used to insert the coding sequence for six histidine residues between the last amino acid codon of the pTrc99-AFFEN1 coding region and the stop codon. The resulting plasmid was termed pTrcAfuHis. The protein was then expressed as described and purified by binding to a Ni⁺⁺ affinity column.

5

Construction of Afu(Y236A), A46-5

Two overlapping PCR fragments were generated from template AfuFEN (SEQ ID NO:556) with primers: 785-73-04 5'-ctg-gtc-ggg-acg-gac-gcc-aat-gag-ggt-gtg-aag-3' (SEQ ID NO: 547) and 700-10-03 5'-ctt-ctc-tca-tcc-gcc-aaa-aca-gcc-3' (SEQ ID NO:343), and primers: 390-076-08 5'-tgt-gga-att-gtg-agc-gg-3' (SEQ ID NO:314) and 785-73-06 5'-gtc-cgt-ccc-gac-cag-aat-3' (SEQ ID NO: 548). The two products were combined and amplified with outside primers 700-10-03 and 390-076-08. The recombinant PCR product was cut with the restriction enzymes NcoI and SalI and ligated into the AfuFEN construct which was prepared by cutting with the same enzymes to yield the Afu(Y236A) construct (DNA sequence SEQ ID NO: 549; amino acid sequence SEQ ID NO: 550).

Construction of Afu(Y236R), A56-9

Two overlapping PCR fragments were generated using the template AfuFEN with primers: 785-73-05 5'-ctg-gtc-ggg-acg-gac-agg-aat-gag-ggt-gtg-aag-3' (SEQ ID NO: 551) and 700-10-03 5'-ctt-ctc-tca-tcc-gcc-aaa-aca-gcc-3' (SEQ ID NO:343), and primers: 390-076-08 5'-tgt-gga-att-gtg-agc-gg-3' (SEQ ID NO:314) and 785-73-06 5'-gtc-cgt-ccc-gac-cag-aat-3' (SEQ ID NO: 548). The two products were combined and amplified with outside primers 700-10-03 and 390-076-08. The recombinant PCR product was cut with the restriction enzymes NcoI and SalI and ligated into the AfuFEN construct which was prepared by cutting with the same enzymes to yield Afu(Y236R) construct (DNA sequence SEQ ID NO: 552; amino acid sequence SEQ ID NO: 553).

2. Chimeras of FEN enzymes and *Thermus* Polymerase derivatives

The following enzyme constructs combine portions of the AfuFEN enzyme polymerase domain and the polymerase domain of *Thermus* polymerases. These combinations were designed based on information generated by molecular modeling.

Construction of Afu336-Tth296(AKK), AT1-1

Two overlapping PCR fragments were generated. The first fragment was made using the AfuFEN construct (SEQ ID NO:556) as template with the primers: 390-076-08 5'-tgt-gga-att-gtg-agc-gg-3' (SEQ ID NO:314) and 390-065-05 5'-gaa-cca-cct-ctc-aag-cgt-gg-3' (SEQ ID NO: 554). The second fragment was made using TthAKK (SEQ ID NO:558) as template and the primers: 700-049-01 5'-acg-ctt-gag-agg-tgg-ttc-ctg-gag-gag-gcc-ccc-tgg-3' (SEQ ID NO: 555) and 390-076-09 5'-taa-tct-gta-tca-ggc-tg-3' (SEQ ID NO:557). The two products contain a region of sequence overlap, and were combined and amplified with outside primers 390-076-08 and 390-076-09 in a recombinant PCR reaction. The recombinant PCR product was cut with the restriction enzymes BspEI and SalI and ligated into the AfuFEN construct which was prepared by cutting with the same enzymes to yield Afu336-Tth296(AKK) construct (DNA sequence SEQ ID NO: 559; amino acid sequence SEQ ID NO: 560).

Construction of Afu328-Tth296(AKK), AT2-3

Two overlapping PCR fragments were generated, the first by primers: 390-076-08 5'-tgt-gga-att-gtg-agc-gg-3' (SEQ ID NO:314) and 700-049-02 5'-ggg-tga-ctt-cag-agc-ttt-gag-3' (SEQ ID NO: 561) with template AfuFEN (DNA sequence SEQ ID NO:556), and the second by primers: 700-049-03 5'-aaa-gct-ctg-aag-tca-acc-ctg-gag-gag-gcc-ccc-tgg-3' (SEQ ID NO: 562) and 390-076-09 5'-taa-tct-gta-tca-ggc-tg-3' (SEQ ID NO:557) with template TthAKK (DNA sequence SEQ ID NO:558). The two products were combined and amplified with outside primers 390-076-08 and 390-076-09. The recombinant PCR product was cut with the restriction enzymes BspEI and SalI and ligated into the vector pAfuFEN which was prepared by cutting with the same enzymes to yield Afu336-Tth296(AKK) construct (DNA sequence SEQ ID NO: 563; amino acid sequence SEQ ID NO: 564).

Construction of Afu336-Taq5M

The Taq5M construct (SEQ ID NO:41) was cut with the enzymes NotI and Sall and the smaller insert fragment was gel isolated. The Afu336-Tth296(AKK) construct (DNA sequence SEQ ID NO: 559) was also cut with the same restriction enzymes and the larger vector fragment was purified. The insert (Taq5M polymerase domain) was then ligated into the vector to generate the Afu336-Taq5M construct (DNA sequence SEQ ID NO: 565; amino acid sequence SEQ ID NO: 566).

Construction of Afu336-TaqDN

The TaqDNHT construct was cut with the enzymes NotI and Sall and the smaller insert fragment was gel isolated. The Afu336-Tth296(AKK) construct (DNA sequence SEQ ID NO: 559) was also cut with the same restriction enzymes and the larger vector fragment was purified. The insert (TaqDN polymerase domain) was then ligated into the vector to generate the Afu336-Taq5M construct (DNA sequence SEQ ID NO: 567; amino acid sequence SEQ ID NO: 568).

Random Chimerization of *Thermus* Polymerases

Numerous enzymes with altered functions were generated via random chimerization of *Thermus* polymerases based on the principal of DNA shuffling (Volkov and Arnold, *Methods in Enzymology* 328:456 [2000]). The procedure below was used to develop the following chimeras.

The genes of interest for random chimerization were PCR amplified with primers: 390-076-08 5'-tgt-gga-att-gtg-agc-gg-3' (SEQ ID NO:314) and 209-074-02: 5'-gtg-gcc-tcc-ata-tgg-gcc-agg-ac-3' (SEQ ID NO:316) to generate the approximately 1.4 kbp templates. About 2 µg of the DNA templates were mixed in equal proportion, and then digested with DNase I (0.33 U) in a 30 µl reaction at 15° C for approximately 1 minute to generate fragments 50-200 bp in size. DNA fragments were purified in a 4 % agarose gel and extracted by QIAEXII gel extraction kit (QIAGEN) according to manufacturer's instructions. The purified fragments (10 µl) were added to 10 µl of 2X PCR pre-mix (5-fold diluted, cloned Pfu buffer, 0.4 mM each dNTP, 0.06 U/µl cloned Pfu DNA polymerase, STRATAGENE Cat.# 600153, with accompanying buffer) for the fragment

reassembly reaction (PCR program: 3 min 94 °C followed by 40 cycles of 30 sec 94 °C, 1 min 52 °C, 2 min + 5 s/cycle 72 °C, followed by 4 min at 72 °C). The reassembled products (1 µl of a 10-fold dilution) were then PCR amplified with a pair of nested primers: 072-090-01 5'-gag-cgg-ata-aca-att-tca-cac-agg-3' (SEQ ID NO: 569) and 189-082-01 5'-tgc-ccg-gtg-cac-gcg-gcc-gcc-cct-gca-ggc-3' (SEQ ID NO: 570) using the CLONTECH GC melt cDNA PCR kit (Cat.# K1907-Y) according to the manufacturer's instructions. The purified PCR products were digested with restriction enzymes EcoRI and NotI and then ligated into the TthAKK construct that was prepared by cutting with the same enzymes. The ligation mixture was transformed into JM109 competent cells and colonies were screened for enzyme activity.

1. Generation of Random Chimeras S26 and S36

The templates used to develop these chimeric enzymes were the nuclease domains from TthAKK, TaqTthAKK, TscTthAKK and the TfiTthAKK constructs. Two clones were found to show improvement of activity based on primary activity screening. Random chimeras S26 (DNA sequence SEQ ID NO: 571; amino acid sequence SEQ ID NO: 572) and S36 (DNA sequence SEQ ID NO: 573; amino acid sequence SEQ ID NO: 574) were then sequenced and isolated.

2. Introduction of L107F/E108T, L109F/V110T and K69E mutations in S26 and S36 to improve substrate specificity

Construction of S26(FT)

Site specific mutagenesis was performed on pS26 DNA using the mutagenic primers: 785-008-03 5'-ttt-acc-cgc-ctc-gag-gtg-ccg-ggc-3' (SEQ ID NO: 489) and 680-21-03 5'-cgg-cac-ctc-gag-gcg-ggt-aaa-gcc-caa-aag-gtc-cac-3' (SEQ ID NO: 490) to introduce the L107F/E108T mutations and generate S26(FT) (DNA sequence SEQ ID NO: 575; amino acid sequence SEQ ID NO: 576).

Construction of S36(FT)

Site specific mutagenesis was performed on pS36 DNA using the mutagenic primers: 785-096-01: 5'-gtg-gac-ctt-ctg-ggc-ttt-acc-cgc-ctc-gag-gcc-ccg-3' (SEQ ID NO:

486) and 785-096-02: 5'-cgg-ggc-ctc-gag-gcg-ggt-aaa-gcc-cag-aag-gtc-cac-3' (SEQ ID NO: 487) to introduce the L109F/V110T mutations and generate S36(FT) (DNA sequence SEQ ID NO: 577; amino acid sequence SEQ ID NO: 578).

Construction of S26(K69E)

Site specific mutagenesis was performed on S26 DNA using the mutagenic primers: 5'-atc-gtg-gtc-ttt-gac-gcc-gag-gcc-ccc-tcc-ttc-c-3' (SEQ ID NO: 492) and 5'-gga-agg-agg-ggg-cct-cgg-cgt-caa-aga-cca-cga-t-3' (SEQ ID NO: 493) to introduce the K69E mutation and to generate S26(K69E) (DNA sequence SEQ ID NO: 579; amino acid sequence SEQ ID NO: 580).

Construction of S26(FT/K69E)

Site specific mutagenesis was performed on S26(FT) DNAs using the mutagenic primers: 5'-atc-gtg-gtc-ttt-gac-gcc-gag-gcc-ccc-tcc-ttc-c-3' (SEQ ID NO: 492) and 5'-gga-agg-agg-ggg-cct-cgg-cgt-caa-aga-cca-cga-t-3' (SEQ ID NO: 493) to introduce the K69E mutation and to generate S26(FT/K69E) (DNA sequence SEQ ID NO: 581; amino acid sequence SEQ ID NO: 582).

3. More random chimeras, N3D7, N1A12 N1C4, and N2C3

The templates used to generate these chimerics were the nuclease domains of Tth(K69E)AKK, Taq(K69E)TthAKK, Tsc(K69E)TthAKK and Tfi(K69E)TthAKK constructs. Four clones were shown to have improved activity based on the primary activity screening. Random chimeras N3D7 (DNA sequence SEQ ID NO: 583; amino acid sequence SEQ ID NO: 584), N1A12 (DNA sequence SEQ ID NO: 585; amino acid sequence SEQ ID NO: 586), N1C4 (DNA sequence SEQ ID NO: 587; amino acid sequence SEQ ID NO: 588) and N2C3 (DNA sequence SEQ ID NO: 589; amino acid sequence SEQ ID NO: 590) were then sequenced and isolated respectively.

EXAMPLE 10

Test For The Dependence Of An Enzyme On

The Presence Of An Upstream Oligonucleotide

When choosing a structure-specific nuclease for use in a sequential invasive cleavage reaction it is preferable that the enzyme have little ability to cleave a probe 1) in the absence of an upstream oligonucleotide, and 2) in the absence of overlap between the

upstream oligonucleotide and the downstream probe oligonucleotide. Figs. 26a-e depicts several structures that can be used to examine the activity of an enzyme is confronted with each of these types of structures. The structure **a** (Fig. 26a) shows the alignment of a probe oligonucleotide with a target site on bacteriophage M13 DNA (M13 sequences shown in Fig. 26 are provided in SEQ ID NO:217) in the absence of an upstream oligonucleotide. Structure **b** (Fig. 26b) is provided with an upstream oligonucleotide that does not contain a region of overlap with the labeled probe (the label is indicated by the star). In structures **c**, **d** and **e** (Figs. 26c-e) the upstream oligonucleotides have overlaps of 1, 3 or 5 nucleotides, respectively, with the downstream probe oligonucleotide and each of these structures represents a suitable invasive cleavage structure. The enzyme *Pfu* FEN-1 was tested for activity on each of these structures and all reactions were performed in duplicate.

Each reaction comprised 1 μ M 5' TET labeled probe oligonucleotide 89-15-1 (SEQ ID NO:212), 50 nM upstream oligonucleotide (either oligo 81-69-2 [SEQ ID NO:216], oligo 81-69-3 [SEQ ID NO:215], oligo 81-69-4 [SEQ ID NO:214], oligo 81-69-5 [SEQ ID NO:213], or no upstream oligonucleotide), 1 fmol M13 target DNA, 10 mg/ml tRNA and 10 ng of *Pfu* FEN-1 in 10 μ l of 10 mM MOPS (pH 7.5), 7.5 mM $MgCl_2$ with 0.05% each of Tween 20 and Nonidet P-40.

All of the components except the enzyme and the $MgCl_2$ were assembled in a final volume of 8 μ l and were overlaid with 10 μ l of CHILLOUT liquid wax. The samples were heated to the reaction temperature of 69°C. The reactions were started by the addition of the *Pfu* FEN-1 and $MgCl_2$, in a 2 μ l volume. After incubation at 69°C for 30 minutes, the reactions were stopped with 10 μ l of 95% formamide, 10 mM EDTA, 0.02% methyl violet. Samples were heated to 90°C for 1 min immediately before electrophoresis through a 20% denaturing acrylamide gel (19:1 cross-linked), with 7 M urea, in a buffer of 45 mM Tris-Borate, pH 8.3, 1.4 mM EDTA. Gels were then analyzed with a FMBIO-100 Hitachi FMBIO fluorescence imager. The resulting image is displayed in Fig. 27.

In Fig. 27, lanes labeled "a" contain the products generated from reactions conducted without an upstream oligonucleotide (structure **a**), lanes labeled "b" contain an upstream oligonucleotide that does not invade the probe/target duplex (structure **b**).

Lanes labeled "c", "d" and "e" contain the products generated from reactions conducted using an upstream oligonucleotide that invades the probe/target duplex by 1, 3 or 5 bases, respectively. The size (in nucleotides) of the uncleaved probe and the cleavage products is indicated to the left of the image in Fig. 27. Cleavage of the probe was not detectable when structures **a** and **b** were utilized. In contrast, cleavage products were generated when invasive cleavage structures were utilized (structures **c-e**). These data show that the *Pfu* FEN-1 enzyme requires an overlapping upstream oligonucleotide for specific cleavage of the probe.

Any enzyme may be examined for its suitability for use in a sequential invasive cleavage reaction by examining the ability of the test enzyme to cleave structures **a-e** (it is understood by those in the art that the specific oligonucleotide sequences shown in Figs. 26a-e need not be employed in the test reactions; these structures are merely illustrative of suitable test structures). Desirable enzymes display little or no cleavage of structures **a** and **b** and display specific cleavage of structures **c-e** (*i.e.*, they generate cleavage products of the size expected from the degree of overlap between the two oligonucleotides employed to form the invasive cleavage structure).

EXAMPLE 11

Use Of The Products Of A First Invasive Cleavage Reaction To Enable A Second Invasive Cleavage Reaction With A Net Gain In Sensitivity

As discussed in the Description of The Invention above, the detection sensitivity of the invasive cleavage reaction can be increased by the performing a second round of invasive cleavage using the products of the first reaction to complete the cleavage structure in the second reaction. In this Example, the use of a probe that, when cleaved in a first invasive cleavage reaction, forms an integrated INVADER oligonucleotide and target molecule for use in a second invasive cleavage reaction, is illustrated.

A first probe was designed to contain some internal complementarity so that when cleaved in a first invasive cleavage reaction the product ("Cut Probe 1") could form a target strand comprising an integral INVADER oligonucleotide. A second probe was provided in the reaction that would be cleaved at the intended site when hybridized to the newly formed target/INVADER molecule. To demonstrate the gain in signal due to the

performance of sequential invasive cleavages, a standard invasive cleavage assay, as described above, was performed in parallel.

All reactions were performed in duplicate. Each standard (*i.e.*, non-sequential) invasive cleavage reaction comprised 1 μ M 5' fluorescein-labeled probe oligo 073-182 (5' FI-AGAAAGGAAGGGAAGAAAGCGAA-3'; SEQ ID NO:591), 10 nM upstream oligo 81-69-4 (5'-CTTGACGGGGAAAGCCGGCGAACGTGGCGA-3'; SEQ ID NO:214), 10 to 100 attomoles of M13 target DNA, 10 mg/ml tRNA and 10 ng of *Pfu* FEN-1 in 10 μ l of 10 mM MOPS (pH 7.5), 8 mM MgCl₂ with 0.05% each of Tween 20 and Nonidet P-40. All of the components except the enzyme and the MgCl₂ were assembled in a volume of 7 μ l and were overlaid with 10 μ l of CHILLOUT liquid wax. The samples were heated to the reaction temperature of 62°C. The reactions were started by the addition of the *Pfu* FEN-1 and MgCl₂, in a 2 μ l volume. After incubation at 62°C for 30 minutes, the reactions were stopped with 10 μ l of 95% formamide, 10 mM EDTA, 0.02% methyl violet.

Each sequential invasive cleavage reaction comprised 1 μ M 5' fluorescein-labeled oligonucleotide 073-191 (the first probe or "Probe 1", 5' FI-TGGAGGTCAAAACATCG ATAAGTCGAAGAAAGGAAGGGAAGAAAT-3'; SEQ ID NO:592), 10 nM upstream oligonucleotide 81-69-4 (5'-CTTGACGGGGAAA GCCGGCGAACGTGGCGA-3'; SEQ ID NO:214), 1 μ M of 5' fluorescein labeled oligonucleotide 106-32 (the second probe or "Probe 2", 5' FI-TGTTTTGACCT CCA-3'; SEQ ID NO:593), 1 to 100 amol of M13 target DNA, 10 mg/ml tRNA and 10 ng of *Pfu* FEN-1 in 10 μ l of 10 mM MOPS (pH 7.5), 8 mM MgCl₂ with 0.05% each of Tween 20 and Nonidet P-40. All of the components except the enzyme and the MgCl₂ were assembled in a volume of 8 μ l and were overlaid with 10 μ l of CHILLOUT liquid wax. The samples were heated to the reaction temperature of 62°C (this temperature is the optimum temperature for annealing of Probe 1 to the first target). The reactions were started by the addition of *Pfu* FEN-1 and MgCl₂, in a 2 μ l volume. After incubation at 62°C for 15 minutes, the temperature was lowered to 58°C (this temperature is the optimum temperature for annealing of Probe 2 to the second target) and the samples were incubated for another 15 min. Reactions

were stopped by the addition of 10 μ l of 95% formamide, 20 mM EDTA, 0.02% methyl violet.

Samples from both the standard and the sequential invasive cleavage reactions were heated to 90° C for 1 min immediately before electrophoresis through a 20% denaturing acrylamide gel (19:1 cross-linked), with 7 M urea, in a buffer of 45 mM Tris-Borate, pH 8.3, 1.4 mM EDTA. The gel was then analyzed with a Molecular Dynamics FluorImager 595. The resulting image is displayed in Fig. 28a. A graph showing measure of fluorescence intensity for each of the product bands is shown in Fig. 28b.

In Fig. 28a, lanes 1-5 contain the products generated in standard invasive cleavage reactions that contained either no target (lane 1), 10 amol of target (lanes 2 and 3) or with 100 amol of target (lanes 4 and 5). The uncleaved probe is seen as a dark band in each lane about half way down the panel and the cleavage products appear as a smaller black band near the bottom of the panel, the position of the cleavage product is indicated by an arrow head to the left of Fig. 28a. The gray ladder of bands seen in lanes 1-5 is due to the thermal degradation of the probe and is not related to the presence or absence of the target DNA. The remaining lanes display products generated in sequential invasive cleavage reactions that contained 1 amol of target (lanes 6 and 7), 10 amol of target (lanes 8 and 9) and 100 amol of target (lanes 10 and 11). The uncleaved first probe (Probe 1; labeled "1 uncut") is seen near the top of the panel, while the cleaved first probe is indicated as "1: cut". Similarly, the uncleaved and cleaved second probe are indicated as "2: uncut" and "2: cut," respectively.

The graph shown in Fig. 28b compares the amount of product generated from the standard reaction ("Series 1") to the amount of product generated from the second step of the sequential reaction ("Series 2"). The level of background fluorescence measured from a reaction that lacked target DNA was subtracted from each measurement. It can be seen from the table located below the graph that the signal from the standard invasive cleavage assay that contained 100 attomoles of target DNA was nearly identical to the signal from the sequential invasive cleavage assay in which 1 attomole of target was present, indicating that the inclusion of a second cleavage structure increases the sensitivity of the assay 100 to 200-fold. This boost in signal allows easy detection of

target nucleic acids at the sub attomole level using the sequential invasive cleavage assay, while the standard assay, when performed using this enzyme for only 30 minutes, does not generate detectable product in the presence of 10 attomoles of target.

When the amount of target was decreased by 10 or 100 fold in the sequential
5 invasive cleavage assay, the intensity of the signal was decreased by the same proportion. This indicates that the quantitative capability of the invasive cleavage assay is retained even when reactions are performed in series, thus providing a nucleic acid detection method that is both sensitive and quantitative.

While in this Example, the two probes used had different optimal hybridization
10 temperatures (*i.e.*, the temperature empiracally determined to give the greatest turnover rate in the given reaction conditions), the probes may also be selected (*i.e.*, designed) to have the same optimal hybridization temperature so that a temperature shift during incubation is not necessary.

15 **EXAMPLE 12**

The Products Of A Completed Sequential Invasive Cleavage Reaction Cannot Cross Contaminate Subsequent Similar Reactions

As discussed in the Description of the Invention, the serial nature of the multiple
invasive cleavage events that occur in the sequential invasive cleavage reaction, in
20 contrast to the reciprocating nature of the polymerase chain reaction and similar doubling assays, means that the sequential invasive cleavage reaction is not subject to contamination by the products of like reactions because the products of the first cleavage reaction do not participate in the generation of new signal in the second cleavage reaction. If a large amount of a completed reaction were to be added to a newly
25 assembled reaction, the background that would be produced would come from the amount of target that was also carried in, combined with the amount of already-cleaved probe that was carried in. In this Example, it is demonstrated that a very large portion of a primary reaction must be introduced into the secondary reaction to create significant signal.

30 A first or primary sequential invasive cleavage reaction was performed as described above using 100 amol of target DNA. A second set of 5 reactions were

assembled as described in Ex. 11 with the exception that portions of the first reaction were introduced and no additional target DNA was included. These secondary reactions were initiated and incubated as described above, and included 0, 0.01, 0.1, 1, or 10% of the first reaction material. A control reaction including 100 amol of target was included in the second set also. The reactions were stopped, resolved by electrophoresis and visualized as described above, and the resulting image is displayed in Fig. 29. The primary probe, uncut second probe and the cut 2nd probe are indicated on the left as "1: cut", 2: uncut" and 2: cut", respectively.

In Fig. 29, lane 1 shows the results of the first reaction with the accumulated product at the bottom of the panel, and lane 2 show a 1:10 dilution of the same reaction, to demonstrate the level of signal that could be expected from that level of contamination, without further amplification. Lanes 3 through 7 show the results of the secondary cleavage reactions that contained 0, 10, 1, 0.1 or 0.01% of the first reaction material added as contaminant, respectively and lane 8 shows a control reaction that had 100 amol of target DNA added to verify the activity of the system in the secondary reaction. The signal level in lane 4 is as would be expected when 10% of the pre-cleaved material is transferred (as in lane 2) and 10% of the transferred target material from the lane 1 reaction is allowed to further amplify. At all levels of further dilution the signal is not readily distinguished from background. These data demonstrate that while a large-scale transfer from one reaction to another may be detectable, cross contamination by the minute quantities that would be expected from aerosol or from equipment contamination would not be easily mistaken for a false positive result. These data also demonstrate that when the products of one reaction are deliberately carried over into a fresh sample, these products do not participate in the new reaction, and thus do not affect the level of target-dependent signal that may be generated in that reaction.

EXAMPLE 13

Detection of Human Cytomegalovirus Viral DNA by Invasive Cleavage

The previous Example demonstrates the ability of the invasive cleavage reaction to detect minute quantities of viral DNA in the presence of human genomic DNA. In this Example, the probe and INVADER oligonucleotides were designed to target the 3104-3061 region of the major immediate early gene of human cytomegalovirus (HCMV) as shown in Fig. 30. In Fig. 30, the INVADER oligo (89-44; SEQ ID NO:219) and the fluorescein (Fl)-labeled probe oligo (89-76; SEQ ID NO:218) are shown annealed along a region of the HCMV genome corresponding to nucleotides 3057-3110 of the viral DNA (SEQ ID NO:220). The probe used in this Example is a poly-pyrimidine probe and as shown herein the use of a poly-pyrimidine probe reduces background signal generated by the thermal breakage of probe oligos.

The genomic viral DNA was purchased from Advanced Biotechnologies, Incorporated (Columbia, MD). The DNA was estimated (but not certified) by personnel at Advanced Biotechnologies to be at a concentration of 170 amol (1×10^8 copies) per microliter. The reactions were performed in quadruplicate. Each reaction comprised 1 μ M 5' fluorescein labeled probe oligonucleotide 89-76 (SEQ ID NO:218), 100 nM INVADER oligonucleotide 89-44 (SEQ ID NO:219), 1 ng/ml human genomic DNA, and one of five concentrations of target HCMV DNA in the amounts indicated above each lane in Fig. 31, and 10 ng of *Pfu* FEN-1 in 10 μ l of 10 mM MOPS (pH 7.5), 6 mM $MgCl_2$ with 0.05% each of Tween 20 and Nonidet P-40. All of the components except the labeled probe, enzyme and $MgCl_2$ were assembled in a final volume of 7 μ l and were overlaid with 10 μ l of CHILLOUT liquid wax. The samples were heated to 95°C for 5 min, then reduced to 62°C. The reactions were started by the addition of probe, *Pfu* FEN-1 and $MgCl_2$, in a 3 μ l volume. After incubation at 62°C for 60 minutes, the reactions were stopped with 10 μ l of 95% formamide, 10 mM EDTA, 0.02% methyl violet. Samples were heated to 90°C for 1 min immediately before electrophoresis through a 20% acrylamide gel (19:1 cross-linked), with 7 M urea, in a buffer of 45 mM Tris-Borate, pH 8.3, 1.4 mM EDTA. Gels were then analyzed with a Molecular Dynamics FluorImager 595.

The resulting image is displayed in Fig. 31. The replicate reactions were run in groups of four lanes with the target HCMV DNA content of the reactions indicated above each set of lanes (0-170 amol). The uncleaved probe is seen in the upper third of the panel ("Uncut 89-76") while the cleavage products are seen in the lower two-thirds of the panel ("Cut 89-76"). It can be seen that the intensity of the accumulated cleavage product is proportional to the amount of the target DNA in the reaction. Furthermore, it can be clearly seen in reactions that did not contain target DNA ("no target") that the probe is not cleaved, even in a background of human genomic DNA. While 10 ng of human genomic DNA was included in each of the reactions shown in Fig. 31, inclusion of genomic DNA up to 200 ng has slight impact on the amount of product accumulated. The data did not suggest that 200 ng per 10 μ l of reaction mixture represented the maximum amount of genomic DNA that could be tolerated without a significant reduction in signal accumulation. For reference, this amount of DNA exceeds what might be found in 0.2 ml of urine (a commonly tested amount for HCMV in neonates) and is equivalent to the amount that would be found in about 5 μ l of whole blood.

These results demonstrate that the standard (*i.e.*, non-sequential) invasive cleavage reaction is a sensitive, specific and reproducible means of detecting viral DNA. Detection of 1.7 amol of target is roughly equivalent to detection of 10^6 copies of the virus. This is equivalent to the number of viral genomes that might be found in 0.2 mls of urine from a congenitally infected neonate (10^2 to 10^6 genome equivalents per 0.2 mls; Stagno *et al.*, J. Infect. Dis., 132:568 [1975]). Use of the sequential invasive cleavage assay would permit detection of even fewer viral DNA molecules, facilitating detection in blood (10^1 to 10^5 viral particles per ml; Pector *et al.*, J. Clin. Microbiol., 30:2359 [1992]), which carries a much larger amount of heterologous DNA.

From the above it is clear that the invention provides reagents and methods to permit the detection and characterization of nucleic acid sequences and variations in nucleic acid sequences. The INVADER-directed cleavage reaction and the sequential INVADER-directed cleavage reaction of the present invention provide ideal direct detection methods that combine the advantages of the direct detection assays (*e.g.*, easy quantification and minimal risk of carry-over contamination) with the specificity provided by a dual or tri oligonucleotide hybridization assay.

As indicated in the Description of the Invention, the use of sequential invasive cleavage reactions can present the problem of residual uncut first, or primary, probe interacting with the secondary target, and either competing with the cut probe for binding, or creating background through low level cleavage of the resulting structure.

5 This is shown diagrammatically in Figs. 32 and 33. In Fig. 32, the reaction depicted makes use of the cleavage product from the first cleavage structure to form an INVADER oligonucleotide for a second cleavage reaction. The structure formed between the secondary target, the secondary probe and the uncut primary probe is depicted in Fig. 32, as the right hand structure shown in step 2a. This structure is recognized and cleaved by
10 the 5' nucleases, albeit very inefficiently (*i.e.*, at less than about 1% in most reaction conditions). Nonetheless, the resulting product is indistinguishable from the specific product, and thus may lead to a false positive result. The same effect can occur when the cleaved primary probe creates and integrated INVADER/target (IT) molecule, as described in Example 11; the formation of the undesirable complex is depicted
15 schematically in Fig. 33, as the right hand structure shown in step 2a.

The improvements provided by the inclusion of ARRESTOR oligonucleotides of various compositions in each of these types of sequential INVADER assays are demonstrated in the following Examples. These ARRESTOR oligonucleotides are configured to bind the residual uncut probe from the first cleavage reaction in the series,
20 thereby increasing the efficacy of and reducing the non-specific background in the subsequent reaction(s).

EXAMPLE 14

"ARRESTOR" Oligonucleotides Improve Sensitivity of Multiple Sequential 25 Invasive Cleavage Assays

In this Example, the effect of including an ARRESTOR oligonucleotide on the generation of signal using the IT probe system 33 is demonstrated. The ARRESTOR oligonucleotide hybridizes to the primary probe, mainly in the portion that recognizes the target nucleic acid during the first cleavage reaction. In addition to examining the effects
30 of adding an ARRESTOR oligonucleotide, the effects of using ARRESTOR oligonucleotides that extended in complementarity different distances into the region of

the primary probe that composes the secondary IT structure were also investigated. These effects were compared in reactions that included the target DNA over a range of concentrations, or that lacked target DNA, in order to demonstrate the level of nonspecific (*i.e.*, not related to target nucleic acid) background in each set of reaction conditions.

The target DNA for these reactions was a fragment that comprised the full length of the hepatitis B genome from strain of serotype adw. This material was created using the polymerase chain reaction from plasmid pAM6 (ATCC #45020D). The PCRs were conducted using a vector-based forward primer, oligo # 156-022-001 (5'-

ggcgaccacaccgctcgt-3'; SEQ ID NO:594) and a reverse primer, oligo #156-022-02 (5'-ccacgatgcgtccggcgtag-3'; SEQ ID NO:595) to amplify the full length of the HBV insert, an amplicon of about 3.2kb. The cycling conditions included a denaturation of the plasmid at 95°C for 5 minutes, followed by 30 cycles of 95°C, 30 seconds; 60°C, 40 seconds; and 72°C, 4 minutes. This was followed by a final extension at 72°C for 10 minutes. The resulting amplicon, termed pAM6#2, was adjusted to 2 M NH₄OAc, and collected by precipitation with isopropanol. After drying *in vacuo*, DNA was dissolved in 10 mM Tris pH 8.0, 0.1 mM EDTA. The concentration was determined by OD₂₆₀ measurement, and by INVADER assay with comparison to a standard of known concentration.

The INVADER reactions were conducted as follows. Five master mixes, termed "A," "B," "C," "D," and "E," were assembled; all mixes contained 12.5 mM MOPS, pH 7.5, 500 fmoles primary INVADER oligo #218-55-05 (SEQ ID NO:596), 10 ng human genomic DNA (Novagen) and 30 ng AfuFEN1 enzyme, for every 8 µl of mix. Mix A contained no added HBV genomic amplicon DNA; mix B contained 600 molecules of HBV genomic amplicon DNA pAM6 #2; mix C contained 6,000 molecules pAM6 #2; mix D contained 60,000 molecules pAM6 #2; and mix E contained 600,000 molecules pAM6 #2. The mixes were aliquotted to the reaction tubes, 8 µl/tube: mix A to tubes 1, 2, 11, 12, 21 and 22; mix B to tubes 3, 4, 13, 14, 23 and 24; mix C to tubes 5, 6, 15, 16, 25 and 26; mix D to tubes 7, 8, 17, 18, 27 and 28; and mix E to tubes 9, 10, 19, 20, 29 and 30. The samples were incubated at 95°C for 4 minutes to denature the HBV genomic amplicon DNA. The reactions were then cooled to 67°C, and 2 µl of a mix containing

37.5 mM MgCl₂ and 2.5 pmoles 218-95-06 (SEQ ID NO:597) for every 2 µl, was added to each sample. The samples were incubated at 67°C for 60 minutes. Three secondary reaction master mixes were prepared, all mixes contained 10 pmoles of secondary probe oligonucleotide #228-48-04 (SEQ ID NO:598) for every 2 µl of mix. Mix 2A contained no additional oligonucleotide, mix 2B contained 5 pmoles “ARRESTOR” oligo # 218-95-03 (SEQ ID NO:599) and mix 2C contained 5 pmoles of “ARRESTOR” oligo # 218-95-01 (SEQ ID NO:600). After the 60 minute incubation at 67°C (the primary reaction described above), 2 µl of the secondary reaction mix was added to each sample: Mix 2A was added to samples #1-10; Mix 2B was added to samples #11-20; and Mix 2C was added to samples #21-30. The temperature was adjusted to 52°C and the samples were incubated for 30 minutes at 52°C. The reactions were then stopped by the addition of 10 µl of a solution of 95% formamide, 5 mM EDTA and 0.02% crystal violet. All samples were heated to 95°C for 2 minutes, and 4 µl of each sample were resolved by electrophoresis through 20% denaturing acrylamide gel (19:1 cross-linked) with 7 M urea, in a buffer containing 45 mM Tris-Borate (pH8.3) and 1.4 mM EDTA. The results were imaged using the Molecular Dynamics Fluoroimager 595, excitation 488, emission 530. The resulting images are shown in Fig. 34.

In Fig. 34, Panel A shows the results of the target titration when no ARRESTOR oligonucleotide was included in the secondary reaction; Panel B shows the results of the same target titration using an ARRESTOR oligonucleotide that extended 2 nt into the non-target complementary region of the primary probe; and Panel C shows the results of the same target titration using an ARRESTOR oligonucleotide that extended 4 nt into the non-target complementary region of the primary probe. The product of the secondary cleavage reaction is seen as a band near the bottom of each panel. The first two lanes of each panel (*i.e.*, 1 and 2, 11 and 12, 21 and 22) lacked target DNA, and the signal that co-migrates with the product band represents the nonspecific background under each set of conditions.

It can be seen by visual inspection of these panels that the background signal is both reduced, and made more predictable, by the inclusion of either species of ARRESTOR oligonucleotide. In addition to reducing the background in the no-target control lanes, the background reduction in the reactions that had the more dilute amounts

of target included is reduced, leading to a signal that is a more accurate reflection of the target contained within the reaction, thus improving the quantitative range of the multiple, sequential invasive cleavage reaction.

To quantify the impact of including the ARRESTOR oligonucleotide in the secondary cleavage reaction under these conditions, the average product band signal from the reactions having the largest amount of target (*i.e.*, averages of the signals from lanes 9 and 10, lanes 19 and 20, and lanes 29 and 30), were compared to the averaged signal from the no-target control lanes for each panel, determine the "fold over background," the factor of signal amplification over background, under each set of conditions. For the reactions without the ARRESTOR oligonucleotide, Panel A, the fold over background was 5.3; for Panel B, the fold over background was 12.7; and for Panel C, the fold over background was 13.4, indicating that in this system inclusion of any ARRESTOR oligonucleotide at least doubled the specificity of the signal over the ARRESTOR oligonucleotide-less reactions, and that the ARRESTOR oligonucleotide that extended slightly farther into the non-target complementary region may be slightly more effective, at least in this embodiment of the system. This clearly shows the benefits of using an ARRESTOR oligonucleotide to enhance the specificity of these reactions, an advantage that is of particular benefit at low levels of target nucleic acid.

EXAMPLE 15

"ARRESTOR" Oligonucleotides Allow use of Higher Concentrations of Primary Probe Without Increasing Background Signal

Increasing the concentration of the probe in the invasive cleavage reaction can dramatically increase the amount of signal generated for a given amount of target DNA. While not intending to limit the explanation to any specific mechanism, this is believed to be caused by the fact that increased concentration of probe increases the rate at which the cleaved probe is supplanted by an uncleaved copy, thereby increasing the apparent turnover rate of the cleavage reaction. Unfortunately, this effect could not heretofore be applied in the primary cleavage reaction of a multiple sequential INVADER assay because the residual uncleaved primary probe can hybridize to the secondary target, in competition with the cleaved molecules, thereby reducing the efficacy of the secondary

reaction. Elevated concentrations of primary probe exacerbate this problem. Further, the resulting complexes, as described above, can be cleaved at a low level, contributing to background. Therefore, increasing the primary probe can have the double negative effect of both slowing the secondary reaction and increasing the level of this form of non target-specific background. The use of an ARRESTOR oligonucleotide to sequester or
5 neutralize the residual primary probe allows this concentration-enhancing effect to be applied to these sequential reactions.

To demonstrate this effect, two sets of reactions were conducted. In the first set of reactions, the reactions were conducted using a range of primary probe concentrations,
10 but no ARRESTOR oligonucleotide was supplied in the secondary reaction. In the second set of reactions, the same probe concentrations were used, but an ARRESTOR oligonucleotide was added for the secondary reactions.

All reactions were performed in duplicate. Primary INVADER reactions were done in a final volume of 10 μ l and contained: 10 mM MOPS, pH 7.5, 7.5 mM $MgCl_2$,
15 500 fm of primary INVADER (218-55-05; SEQ ID NO:596); 30 ng of AfuFEN1 enzyme and 10 ng of human genomic DNA. 100 zeptomoles of HBV pAM6 #2 amplicon was included in all even numbered reactions (by reference to Figs. 35A and B). Reactions included 10 pmoles, 20 pmoles, 50 pmoles, 100 pmoles or 150 pmoles of primary probe (218-55-02; SEQ ID NO:601). MOPS, target and INVADER oligonucleotides were
20 combined to a final volume of 7 μ l. Samples were heat denatured at 95°C for 5 minutes, then cooled to 67°C. During the 5 minute denaturation, $MgCl_2$, probe and enzyme were combined. The primary INVADER reactions were initiated by the addition of 3 μ l of $MgCl_2$, probe and enzyme mix, to the final concentrations indicated above. Reactions were incubated for 30 minutes at 67°C. The reactions were then cooled to 52°C, and
25 each primary INVADER reaction received the following secondary reaction components in a total volume of 4 μ l: 2.5 pmoles secondary target (oligo number 218-95-04; SEQ ID NO:602); 10 pmoles secondary probe (oligo number 228-48-04; SEQ ID NO:598). The reactions that included the ARRESTOR oligonucleotide had either 40 pmoles, 80 pmoles, 200 pmoles, 400 pmoles or 600 pmoles of ARRESTOR oligonucleotide (oligo number
30 218-95-01; SEQ ID NO:600), added at a 4-fold molar excess over the primary probe amount for each reaction, with this mix. Reactions were then incubated at 52°C for 30

minutes. The reactions were stopped by the addition of 10 μ l of a solution of 95% formamide, 10 mM EDTA and 0.02% crystal violet. All samples were heated to 95°C for 1 minute, and 4 μ l of each sample were resolved by electrophoresis through 20% denaturing acrylamide gel (19:1 cross-linked) with 7 M urea, in a buffer containing 45 mM Tris-Borate (pH8.3) and 1.4 mM EDTA. The results were imaged using the Molecular Dynamics Fluoroimager 595, excitation 488, emission 530. The resulting images for the reactions either without or with an ARRESTOR oligonucleotide are shown in Figs. 35A and 35B, respectively. The products of cleavage of the secondary probe are seen as a band near the bottom of each panel.

In Fig. 35A, lane sets 1 and 2 show results with 10 pmoles of primary probe; 3 and 4 had 20 pmoles; 5 and 6 had 50 pmoles; 7 and 8 had 100 pmoles; and 9 and 10 had 150 pmoles. It can be seen by visual examination, that the increases in the amount of primary probe have the combined effect of slightly increasing the background in the no-target lanes (odd numbers) while reducing the specific signal in the presence of target (even numbered lanes), and therefore the reducing the specificity of the reaction if viewed as the measure of "fold over background," demonstrating that the approach of increasing signal by increasing probe cannot be applied in these sequential reactions.

In Fig. 35B, lane sets 1 and 2 show results with 10 pmoles of primary probe; while 3 and 4 had 20 pmoles; 5 and 6 had 50 pmoles; 7 and 8 had 100 pmoles; and 9 and 10 had 150 pmoles. In addition, each reaction included 4-fold molar excess of the ARRESTOR oligonucleotide added before the secondary cleavage reaction. It can be seen by visual examination that the background in the no-target lanes (odd numbers) is lower in all cases, while the specific signal in the presence of target (even numbered lanes) increases with increased amounts of primary probe, leading to a greater "fold over background" sensitivity at this target level.

To quantitatively compare these effects, the fluorescence signal from the products of both non-specific and specific cleavage were measured. The results are depicted graphically in Fig. 35C, graphed as a measure of the percentage of the secondary probe cleaved during the reaction, compared to the amount of primary probe used.

Examination of the plots from the no-target reactions confirms that the background in the absence of the ARRESTOR oligonucleotide is, in general, roughly two-fold higher, and

that both increase slightly with the increasing probe amounts. The specific signals however, diverge between the two sets of reaction more dramatically. While the signal in the no-ARRESTOR oligonucleotide reactions decreases steadily as primary probe was increased, the signal in the ARRESTOR oligonucleotide reactions continued to increase. At the highest primary probe concentrations tested, the no-ARRESTOR oligonucleotide reactions had specific signal that was only 1.7 fold over background, while the ARRESTOR oligonucleotide reactions detected the 100 zmoles (60,000 copies) of target with a signal 6.5 fold over background, thus demonstrating the improvement in the sequential invasive cleavage reaction when an ARRESTOR oligonucleotide is included.

EXAMPLE 16

Modified Backbones Improve Performance of ARRESTOR Oligonucleotides All Natural "ARRESTOR" Oligo With No 3'-Amine

The reactions described in the previous two Examples used ARRESTOR oligonucleotides that were constructed using 2' O-methyl ribose backbone, and that included a positively charged amine group on the 3' terminal nucleotide. The modifications were made specifically to reduce enzyme interaction with the primary probe/ARRESTOR oligonucleotide complex. During the development of the present invention, it was determined that the 2' O-methyl modified oligonucleotides are somewhat resistant to cleavage by the 5' nucleases, just as they are slowly degraded by nucleases when used in antisense applications (*See e.g., Kawasaki et al., J. Med. Chem., 36:831 [1993]*).

The presence of an amino group on the 3' end of an oligonucleotide reduces its ability to direct invasive cleavage. To reduce the possibility that the ARRESTOR oligonucleotide would form a cleavage structure in this way, an amino group was included in the design of the experiments described in this and other Examples.

Initial designs of the ARRESTOR oligonucleotides (sometimes referred to as "blockers") did not include these modifications, and these molecules were found to provide no benefit in reducing background cleavage in the sequential invasive cleavage assay and, in fact, sometimes contributed to background by inducing cleavage at an unanticipated site, presumably by providing some element to an alternative cleavage

structure. The effects of natural and modified ARRESTOR oligonucleotides on the background noise in these reactions are examined in this Example.

The efficacy of an "all-natural ARRESTOR oligonucleotide (*i.e.*, an ARRESTOR oligonucleotide that did not contain any base analogs or modifications) was examined by comparison to an identical reactions that lacked ARRESTOR oligonucleotide. All reactions were performed in duplicate, and were conducted as follows. Two master mixes were assembled, each containing 12.5 mM MOPS, pH 7.5, 500 fmoles primary INVADER oligonucleotide #218-55-05 (SEQ ID NO:596), 10 ng human genomic DNA (Novagen) and 30 ng AfuFEN1 enzyme for every 8 μ l of mix. Mix A contained no added HBV genomic amplicon DNA, mix B contained 600,000 molecules of HBV genomic amplicon DNA, pAM6 #2. The mixes were distributed to the reaction tubes, in aliquots of 8 μ l/tube as follows: mix A to tubes 1, 2, 5 and 6; and mix B to tubes 3, 4, 7 and 8. The samples were incubated at 95°C for 4 minutes to denature the HBV genomic amplicon DNA. The reactions were then cooled to 67°C and 2ul of a mix containing 37.5 mM MgCl₂ and 10 pmoles 218-55-02B (SEQ ID NO:603) for every 2 μ l, was added to each sample. The samples were then incubated at 67°C for 30 minutes. Two secondary reaction master mixes were prepared, each containing 10 pmoles of secondary probe oligo #228-48-04N (SEQ ID NO:604) and 2.5 pmoles of secondary target oligonucleotide #218-95-04 (SEQ ID NO:602) for every 3 μ l of mix. Mix 2A contained no additional oligonucleotide, while mix 2B contained 50 pmoles of the natural "ARRESTOR" oligonucleotide #241-62-02 (SEQ ID NO:605). After the initial 30 minute incubation at 67°C, the temperature was adjusted to 52°C, and 3 μ l of a secondary reaction mix was added to each sample, as follows: Mix 2A was added to samples #1-4; and Mix 2B was added to samples #5-8. The samples were then incubated for 30 minutes at 52°C. The reactions were then stopped by the addition of 10 μ l of a solution of 95% formamide, 10 mM EDTA and 0.02% crystal violet.

All of the samples were heated to 95°C for 2 minutes, and 4 μ l of each sample were resolved by electrophoresis through a 20% denaturing acrylamide gel (19:1 cross-linked) with 7 M urea, in a buffer containing 45 mM Tris-Borate (pH8.3) and 1.4 mM

EDTA. The results were imaged using the Molecular Dynamics Fluoroimager 595, excitation 488, emission 530. The resulting image is shown in Fig. 36A.

To compare the effects of the various modifications made to the ARRESTOR oligonucleotides, reactions were performed using ARRESTOR oligonucleotides having all natural bases, but including a 3' terminal amine; ARRESTOR oligonucleotides having a 3' portion composed of 2' O-methyl nucleotides, plus the 3' terminal amine; and ARRESTOR oligonucleotides composed entirely of 2' O-methyl nucleotides, plus the 3' terminal amine. These were compared to reactions performed without an ARRESTOR oligonucleotide. The reactions were conducted as follows. Two master mixes were assembled, all mixes contained 14.3 mM MOPS, pH 7.5, 500 fmoles primary INVADER oligo #218-55-05 (SEQ ID NO:596) and 10 ng human genomic DNA (Novagen) for every 7 μ l of mix. Mix A contained no added HBV genomic amplicon DNA, mix B contained 600,000 molecules of HBV genomic amplicon DNA, pAM6 #2. The mixes were distributed to the reaction tubes, at 7 μ l/tube: mix A to tubes 1, 2, 5, 6, 9, 10, 13 and 14; and mix B to tubes 3, 4, 7, 8, 11, 12, 15 and 16. The samples were warmed to 95°C for 4 minutes to denature the HBV DNA. The reactions were then cooled to 67°C and 3 μ l of a mix containing 25 mM MgCl₂, 25 pmoles 218-55-02B (SEQ ID NO:603) and 30 ng AfuFEN1 enzyme per 3 μ l, were added to each sample. The samples were then incubated at 67°C for 30 minutes. Four secondary reaction master mixes were prepared; all mixes contained 10 pmoles of secondary probe oligonucleotide #228-48-04B (SEQ ID NO:606) and 2.5 pmoles of secondary target oligonucleotide #218-95-04 (SEQ ID NO:602) for every 3 μ l of mix. Mix 2A contained no additional oligonucleotide, while mix 2B contained 100 pmoles of the natural+amine ARRESTOR oligonucleotide # 241-62-01 (SEQ ID NO:607), mix 2C contained 100 pmoles of partially O-methyl+amine oligonucleotide # 241-62-03 (SEQ ID NO:608) and mix 2D contained 100 pmoles of all O-methyl+amine oligonucleotide # 241-64-01 (SEQ ID NO:609). After the initial 30 minute incubation at 67°C, the temperature was adjusted to 52°C and 3 μ l of a secondary reaction mix was added to each sample, as follows: mix 2A was added to samples #1-4; mix 2B was added to samples #5-8; mix 2C was added to samples #9-12; and mix 2D was added to samples #13-16. The samples were incubated for 30 minutes at 52°C, then

stopped by the addition of 10 μ l of a solution of 95% formamide, 10 mM NaEDTA, and 0.2% crystal violet.

All samples were heated to 95°C for 2 minutes, and 4 μ l of each sample were resolved by electrophoresis through a 20% denaturing acrylamide gel (19:1 cross-linked) with 7 M urea, in a buffer containing 45 mM Tris-Borate (pH8.3) and 1.4 mM EDTA. The results were imaged using the Molecular Dynamics Fluoroimager 595, excitation 488, emission 530. The resulting image is shown in Fig. 36B.

In Fig. 36A, the left-hand panel shows the reactions that lacked an ARRESTOR oligonucleotide, while the right hand panel shows the data from reactions that included the all natural ARRESTOR oligonucleotide. The first two lanes of each panel are from no-target controls, the second set of lanes contained target. The products of cleavage are visible in the bottom one-fourth of each panel. The position at which the specific reaction products should run is indicated by arrows on left and right.

It can be seen by examination of these data, that the reactions run in the absence of ARRESTOR oligonucleotide show reproducible quality between the replicates, and show significant cleavage only when target is present. In contrast, the addition of another unmodified oligonucleotide into the reactions causes great variation between the replicate lanes (*e.g.*, lanes 5 and 6 were provided with the same reactants, but produced markedly different results). The introduction of the all-natural ARRESTOR oligonucleotide produced, rather than reduced, background in these no-target lanes, and increased cleavage at other sites (*i.e.*, the bands other than those indicated by the arrows flanking the panels). For these reasons the modifications that are described above, the effects of which are shown on Fig. 36B, were incorporated.

The first 4 lanes of Fig. 36B show the products of duplicate reactions without an ARRESTOR oligonucleotide, plus or minus the HBV target (lanes 1, 2, and lanes 3, 4, respectively); The next 4 lanes, 5, 6 and 7, 8 used a natural ARRESTOR oligonucleotide having a 3' terminal amine; lanes 9, 10 and 11, 12 used the ARRESTOR oligonucleotide with a 3' portion composed of 2' O-methyl nucleotides, and having a 3' terminal amine; lanes 13, 14 and 15, 16 used the ARRESTOR oligonucleotide composed entirely of 2' O-methyl nucleotides and having a 3' terminal amine. The products of cleavage of the secondary probe are visible in the lower one third of each panel.

Visual inspection of these data shows that the addition of the 3' terminal amine to the natural ARRESTOR oligonucleotide suppresses the aberrant cleavage seen in Fig. 36A, but this ARRESTOR oligonucleotide does not improve the performance of the reaction, as compared to the no-ARRESTOR oligonucleotide controls. In contrast, the use of the 2' O-methyl nucleotides in the body of the ARRESTOR oligonucleotide does reduce background, whether partially or completely substituted. To quantify the relative effects of these modifications, the fluorescence from each of the co-migrating product bands was measured, the signals from the duplicate lanes were averaged and the "fold over background" was calculated for each reaction containing target nucleic acid.

When ARRESTOR oligonucleotide was omitted, the target-specific signal (lanes 3, 4) was 27-fold over the no target background; the natural ARRESTOR oligonucleotide +amine gave a signal of 17-fold over background; the partial 2' O-methyl+ amine gave a signal of 47-fold over background; and the completely 2' O-methyl+ amine gave a signal of 33 fold over background.

These Figures show that both modifications can have a beneficial effect on the specificity of the multiple, sequential invasive cleavage assay. They also show that the use of the 2' O-methyl substituted backbone, either partial or entire, markedly improves the specificity of these reactions. It is intended that, in various embodiments of the present invention, any number of modifications that make either the ARRESTOR oligonucleotide or the complex it forms with the primary target resistant to nucleases will provide similar enhancement.

EXAMPLE 17

Effect of ARRESTOR Oligonucleotide Length on Signal Enhancement in Multiple Sequential Invasive Cleavage Assays

As noted in the Description of the Invention, the optimal length for an ARRESTOR oligonucleotide depends upon the design of the other nucleic acid elements of the INVADER reaction, particularly on the design of the primary probe. In this Example, the effects of varying the length of the ARRESTOR oligonucleotide were explored in systems using two different secondary probes. A schematic diagram showing these ARRESTOR oligonucleotides aligned as they would hybridize to the primary probe

oligonucleotide is provided in Fig. 37C. In this Figure, the region of the primary probe that recognizes the target nucleic acid is shown underlined; the non-underlined portion, plus the first underlined base is the portion that is released by the first cleavage, and goes on to participate in the second or subsequent cleavage structure.

5 All reactions were performed in duplicate. The INVADER reactions were done in a final volume of 10 μ l final volume containing 10 mM MOPS, pH 7.5, mM $MgCl_2$, 500 fmoles of primary INVADER 241-95-01, (SEQ ID NO:610), 25 pmoles of primary probe 241-95-02 (SEQ ID NO:611), 30 ng of AfuFEN1 enzyme, and 10 ng of human genomic DNA, and if included, 1 amoles of HBV amplicon pAM 6 #2. MOPS, target
10 DNA, and INVADER oligonucleotides were combined to a final volume of 7 μ l.

Samples were heat denatured at 95°C for 5 minutes, then cooled to 67°C. During the 5 minute denaturation, $MgCl_2$, probe and enzyme were combined. The primary INVADER reactions were initiated by the addition of 3 μ l of $MgCl_2$, probe and enzyme mix, to the final concentrations indicated above. Reactions were incubated for 30 minutes at 67°C.

15 The reaction were then cooled to 52°C, and each primary INVADER reaction received the following secondary reaction components in a total volume of 3 μ l: 2.5 pmoles secondary target 241-95-07 (SEQ ID NO:612), 10 pmoles of either secondary probe 228-48-04 (SEQ ID NO:598), or 228-48-04N (SEQ ID NO:604) and 100 pmoles of an ARRESTOR oligonucleotide, either 241-95-03 (SEQ ID NO:613), 241-95-04 (SEQ ID
20 NO:614), 241-95-05 (SEQ ID NO:615) or 241-95-06 (SEQ ID NO:616). The ARRESTOR oligonucleotides were omitted from some reactions as controls for ARRESTOR oligonucleotide effects.

The reactions were incubated at 52°C for 34 minutes, and were then stopped by the addition of 10 μ l of 95% formamide, 10 mM EDTA, and 0.02% crystal violet. All
25 samples were heated to 95°C for 1 minute, and 4 μ l of each sample were resolved by electrophoresis through 20% denaturing acrylamide gel (19:1 cross-linked) with 7 M urea, in a buffer containing 45 mM Tris-Borate (pH8.3) and 1.4 mM EDTA. The results were imaged using the Molecular Dynamics Fluoroimager 595, excitation 488, emission 530. The resulting images for the reactions with the shorter and longer secondary probes
30 are shown in Figs. 37A and 37B, respectively.

In each Figure, the products of cleavage are visible as bands in the bottom half of each lane. The first 4 lanes of each Figure show the products of duplicate reactions without an ARRESTOR oligonucleotide, plus or minus the HBV target (lanes sets 1 and 2 respectively); in the next 4 lanes, sets 3 and 4 used the shortest ARRESTOR 241-95-03 (SEQ ID NO:613); lanes 5 and 6 used 241-95-04 (SEQ ID NO:614); lanes 7 and 8 used 241-95-05 (SEQ ID NO:615); and lanes 9 and 10 used 241-95-06 (SEQ ID NO:616).

The principal background of concern is the band that appears in the "no target" control lanes (odd numbers; this band co-migrates with the target-specific signal near the bottom of each gel panel). Visual inspection shows that the shortest ARRESTOR oligonucleotide was the least effective at suppressing this background, and that the efficacy was increased when the ARRESTOR oligonucleotide extended further into the portion that participates in the subsequent cleavage reaction. Even with this difference in effect, it can be seen from these data that there is much latitude in the design of the ARRESTOR oligonucleotide. The choice of lengths will be influenced by the temperature at which the reaction making use of the ARRESTOR oligonucleotide is performed, the lengths of the duplexes formed between the primary probe and the target, the primary probe and the secondary target, and the relative concentrations of the different nucleic acid species in the reactions.

EXAMPLE 18

Effect of ARRESTOR Oligonucleotide Concentration on Signal Enhancement in Multiple Sequential Invasive Cleavage Assays

In examining the effects of including ARRESTOR oligonucleotides in these cleavage reactions, it was of interest to determine if the concentration of the ARRESTOR oligonucleotide in excess of the primary probe concentration would have an effect on yields of either non-specific or specific signal, and if the length of the ARRESTOR oligonucleotide would be a factor. These two variables were investigated in the following Example.

All reactions were performed in duplicate. The primary INVADER reactions were done in a final volume of 10 μ l and contained 10 mM MOPS, pH 7.5; 7.5 mM $MgCl_2$, 500 fmoles of primary INVADER 241-95-01 (SEQ ID NO:610), 25 pmoles of

primary probe 241-95-02 (SEQ ID NO:611), 30 ng of AfuFEN1 enzyme, and 10 ng of human genomic DNA. Where included, the target DNA was 1 amole of HBV amplicon pAM 6 #2, as described above. MOPS, target and INVADER were combined to a final volume of 7 μ l. The samples were heat denatured at 95°C for 5 minutes, then cooled to 67°C. During the 5 minute denaturation, MgCl₂, probe and enzyme were combined. The primary INVADER reactions were initiated by the addition of 3 μ l of MgCl₂, probe and enzyme mix. The reactions were incubated for 30 minutes at 67°C. The reactions were then cooled to 52°C and each primary INVADER reaction received the following secondary reaction components: 2.5 pmoles secondary target 241-95-07 (SEQ ID NO:612), 10 pmoles secondary probe 228-48-04 (SEQ ID NO:598); and, if included, 50, 100 or 200 pmoles of either ARRESTOR oligonucleotide 241-95-03 (SEQ ID NO:613) or 241-95-05 (SEQ ID NO:615), in a total volume of 3 μ l. Reactions were then incubated at 52°C for 35 minutes. Reactions were stopped by the addition of 10 μ l of 95% formamide, 10 mM EDTA, and 0.02% crystal violet. All of the samples were heated to 95°C for 1 minute, and 4 μ l of each sample were resolved by electrophoresis through 20% denaturing acrylamide gel (19:1 cross-linked) with 7 M urea, in a buffer containing 45 mM Tris-Borate (pH8.3), and 1.4 mM EDTA. The results were imaged using the Molecular Dynamics Fluoroimager 595, excitation 488, emission 530. The resulting images are shown as a composite image in Fig. 38.

Each of the duplicate reactions were loaded on the gel in adjacent lanes and are labeled with a single lane number. All odd numbered lanes were no-target controls. Lanes 1 and 2 had no ARRESTOR oligonucleotide added; lanes 3-8 show results from reactions containing the shorter ARRESTOR oligonucleotide, 241-95-03 (SEQ ID NO:613); lanes 9-14 show results from reactions containing the longer ARRESTOR oligonucleotide, 241-95-05 (SEQ ID NO:615). The products of cleavage from the secondary reaction are visible in the bottom one third of each panel. Visual inspection of these data (*i.e.*, comparison of the specific products to the background bands) shows that both ARRESTOR oligonucleotides have some beneficial effect at all concentration.

To quantify the relative effects of ARRESTOR oligonucleotide length and concentration, the fluorescence from each of the co-migrating product bands was measured, the signals from the duplicate lanes were averaged and the "fold over

background" (signal+target/signal-target) was calculated for each reaction containing target nucleic acid. The reaction lacking an ARRESTOR oligonucleotide yielded a signal approximately 27-fold over background. Inclusion of the shorter ARRESTOR oligonucleotide at 50, 100 or 200 pmoles produced products at 42, 51 and 60-fold over background, respectively. This shows that while the short ARRESTOR at the lowest concentration seems to be less effective than the longer ARRESTOR oligonucleotides (See, *e.g.*, previous Example) this can be compensated for by increasing the concentration of ARRESTOR oligonucleotide, and thereby the ARRESTOR oligonucleotide:primary probe ratio.

In contrast, inclusion of the longer ARRESTOR oligonucleotide at 50, 100 or 200 pmoles produced products at 60, 32 and 24 fold over background, respectively. At the lowest concentration, the efficacy of this longer ARRESTOR oligonucleotide relative to the shorter ARRESTOR oligonucleotide is consistent with the previous Example. Increasing the concentration, however, decreased the yield of specific product, suggesting a competition effect with some element of the secondary cleavage reaction.

These data show that the ARRESTOR oligonucleotides can be used to advantage in a number of specific reaction designs. The choice of concentration will be influenced by the temperature at which the reaction making use of the ARRESTOR oligonucleotide is performed, the lengths of the duplexes formed between the primary probe and the target, the primary probe and the secondary target, and between the primary probe and the ARRESTOR oligonucleotide.

Selection of oligonucleotides for target nucleic acids other than the HBV shown here, (*e.g.*, oligonucleotide composition and length), and the optimization of cleavage reaction conditions in accord with the models provided here follow routine methods and common practice well known to those skilled in the methods of molecular biology.

Example 10 demonstrated that some enzymes require an overlap between an upstream INVADER oligonucleotide and a downstream probe oligonucleotide to create a cleavage structure (Figure 27). It has also been determined that the 3' terminal nucleotide of the INVADER oligonucleotide need not be complementary to the target strand, even if it is the only overlapping base in the INVADER oligonucleotide (*e.g.*, as with the HCMV probes shown in Figure 30). The requirement for an overlap can serve as a convenient

basis for detecting single base polymorphisms (SNPs) or mutations in a nucleic acid sample.

For detection of single base variations, at least two oligonucleotides (e.g., a probe and an INVADER oligonucleotide) hybridize in tandem to the target nucleic acid to form the overlapping structure recognized by the CLEAVASE enzyme to be used in the reaction. An unpaired “flap” is included on the 5' end of the probe. The enzyme recognizes the overlap and cleaves off the unpaired flap, releasing it as a target-specific product. Enzymes that have a strong preference for an overlapping structure, i.e., that cleave the overlapping structure at a much greater rate than they cleave a non-overlapping structure include the FEN-1 enzymes from *Archaeoglobus fulgidus* and *Pyrococcus furiosus* and such enzymes are particularly preferred in the detection of mutations and SNPs. .

EXAMPLE 19

Kits for performing the mRNA INVADER Assay

In some embodiments, the present invention provides kits comprising one or more of the components necessary for practicing the present invention. For example, the present invention provides kits for storing or delivering the enzymes of the present invention and/or the reaction components necessary to practice a cleavage assay (e.g., the INVADER assay). By way of example, and not intending to limit the kits of the present invention to any particular configuration or combination of components, the following section describes one embodiment of a kit for practicing the present invention:

In some embodiments, the kits of the present invention provide the following reagents:

| | |
|---|---------------------------|
| CLEAVASE enzyme (e.g., TthAKK) | Primary Oligos |
| RNA Primary Buffer 1 | Secondary Oligos |
| RNA Secondary Buffer 1 | RNA Standard [100amol/μl] |
| tRNA Carrier [20 ng/μl] | 10X Cell Lysis Buffer 1 |
| T ₁₀ e _{0.1} Buffer [10mM Tris•HCl, pH 8, 0.1mM EDTA] | |

Examples of Primary Oligonucleotides and Secondary Oligonucleotides suitable for use with the methods of the present invention are provided in Figure 41. While the oligonucleotides shown therein may find use in a number of the methods, and variations of the methods, of the present invention, these INVADER assay oligonucleotide sets find particular use with kits of the present invention. The oligonucleotide sets shown in Figure 41 may be used as individual sets to detect individual target RNAs, or may be comined in biplex or multiplex reactions for the detection of two or more analytes or controls in a single reaction. It is contemplated that the designs of these probes sets (e.g., the oligonucleotides and/or their sequences) may be adapted for use in DNA detection assays, using the guidelines for reaction design and optimization provided herein. Additional oligonucleotides that find use in detection assays and kits of the present invention are provided in Figure 47.

In some embodiments, a kit of the present invention provides a list of additional components (e.g., reagents, supplies, and/or equipment) to be supplied by a user in order to perform the methods of the invention. For example, and without intending to limit such additional components lists to any particular components, one embodiment of such a list comprises the following:

- RNase-free (e.g., DEPC-treated) H₂O
- Clear CHILLOUT-14 liquid wax (MJ Research) or RNase-free, optical grade mineral oil (Sigma, Cat. No. M-5904)
- Phosphate-buffered saline (no MgCl₂, no CaCl₂)
- 96-well polypropylene microplate (MJ Research, Cat. No. MSP-9601)
- 0.2-ml thin-wall tubes
- Thermaseal well tape (e.g., GeneMate, Cat. No. T-2417-5)
- Multichannel pipets (0.5–10µl, 2.5–20µl, 20–200µl)
- Thermal cycler or other heat source (e.g., lab oven or heating block).
- Fluorescence microplate reader (a preferred plate reader is top-reading, equipped with light filters have the following characteristics:

| Excitation (Wavelength/Bandwidth) | Emission (Wavelength/Bandwidth) |
|--------------------------------------|------------------------------------|
| 485 nm/20 nm | 530 nm/25 nm |
| 560 nm/20 nm | 620 nm/40 nm |

In some embodiments, a kit of the present invention provides a list of optional components (e.g., reagents, supplies, and/or equipment) to be supplied by a user to facilitate performance of the methods of the invention. For example, and without
 5 intending to limit such optional components lists to any particular components, one embodiment of such a list comprises the following:

- tRNA Solution, 20ng/μl (Sigma, R-5636)
- 1X Stop Solution (10mM Tris•HCl, pH 8, 10mM EDTA)
- Black opaque, 96-well microplate (*e.g.*, COSTAR, Cat. No. 3915)
- 10 ▪ Electronic repeat pipet (250μl)

In some embodiments of a kit, detailed protocols are provided. In preferred
 embodiments, protocols for the assembly of INVADER assay reactions (*e.g.*,
 formulations and preferred procedures for making reaction mixtures) are provided. In
 15 particularly preferred embodiments, protocols for assembly of reaction mixtures include
 computational or graphical aids to reduce risk of error in the performance of the methods
 of the present invention (*e.g.*, tables to facilitate calculation of volumes of reagents
 needed for multiple reactions, and plate-layout guides to assist in configuring multi-well
 assay plates to contain numerous assay reactions). By way of example, and without
 20 intending to limit such protocols to any particular content or format, kits of the present
 invention may comprise the following protocol:

I. DETAILED mRNA INVADER ASSAY PROTOCOL

1. Plan the microplate layout for each experimental run. An example microplate layout
 25 for 40 samples, 6 standards, and a No Target Control is shown in Fig. 40. Inclusion of
 a No Target Control (tRNA Carrier or 1X Cell Lysis Buffer 1) and quantitation

standards are required for absolute quantitation.

2. Prepare the Primary Reaction Mix for either the single or biplex assay format. To calculate the volumes of reaction components needed for the assay (X Volume), multiply the number of reactions (for both samples and controls) by 1.25 [**X Volume** (μl) = # reactions x 1.25]. Vortex the Primary Reaction Mix briefly after the last reagent addition to mix thoroughly. Aliquot 5 μl of the Primary Reaction Mix per microplate well (an electronic repeat pipet is recommended for this step).

Primary Reaction Mix

Single Assay Format

| Reaction Components | 1X Volume | <u> </u> X Volume |
|---|------------------------------------|----------------------|
| RNA Primary Buffer 1 | 4.0 μl | |
| <u> </u> Primary Oligos | 0.25 μl | |
| T ₁₀ e _{0.1} Buffer | 0.25 μl | |
| CLEAVASE enz. enzyme | 0.5 μl | |
| Total Mix Volume (1X) | 5.0μl | |

Biplex Assay Format

| Reaction Components | 1X Volume | <u> </u> X Volume |
|---|------------------------------------|----------------------|
| RNA Primary Buffer 1 | 4.0 μl | |
| <u> </u> Primary Oligos | 0.25 μl | |
| <u> </u> Housekeeping Primary Oligos | 0.25 μl | |
| CLEAVASE enzyme | 0.5 μl | |
| Total Mix Volume (1X) | 5.0μl | |

3. Add 5 μl of each No Target Control, standard, or sample (total RNA or cell lysate) to the appropriate well and mix by pipetting up and down 1–2 times. Overlay each reaction with 10 μl of clear CHILLOUT or mineral oil. Seal microplate with Thermaseal well tape.
4. Incubate reactions for 90 minutes at 60°C in a thermal cycler or oven.
5. While the primary reaction is incubating, prepare the Secondary FRET Reaction Mix for the single or biplex format. Calculate the component volumes required (X Volume) by multiplying the number of reactions (for both samples and controls) by

1.25 [X Volume (μl) = # reactions x 1.25 (μl)]. Aliquot the Secondary FRET Reaction Mix into multiple 0.2-ml thin-wall tubes or an 8-well strip (70μl/tube is sufficient for a row of 12 reactions).

5 Secondary FRET Reaction Mix

Single Assay Format

| Reaction Components | 1X Volume | ___X Volume |
|---|--------------|-------------|
| RNA Secondary Buffer 1 | 2.0μl | |
| _____ Secondary Oligos | 1.5μl | |
| T ₁₀ e _{0.1} Buffer | 1.5μl | |
| Total Mix Volume (1X) | 5.0μl | |

Biplex Assay Format

| Reaction Components | 1X Volume | ___X Volume |
|------------------------------|--------------|-------------|
| RNA Secondary Buffer 1 | 2.0μl | |
| _____ Secondary Oligos | 1.5μl | |
| _____ Housekeeping Secondary | 1.5μl | |
| Total Mix Volume (1X) | 5.0μl | |

10

- After the primary reaction incubation is completed, remove the microplate seal, and add 5μl Secondary FRET Reaction Mix per well using a multichannel pipet. Mix by pipetting up and down 1–2 times. Reseal the microplate with the well tape and incubate the microplate at 60°C for 60 or 90 minutes, as indicated in each Product Information Sheet. The secondary reaction incubation time can be varied. See sections 2 of the **PROCEDURAL NOTES FOR OPERATION OF THE mRNA INVADER ASSAY** for details.

15

- Reactions can be read using one of two procedures: Direct Read or Stop and Transfer.

20

NOTE: Remove the microplate seal before reading the microplate.

Direct Read Procedure

- This procedure enables collection of multiple data sets to extend the assay's dynamic range. During the secondary INVADER reaction, read the microplate directly in a top-reading fluorescence microplate reader.
- Recommended settings for a PerSeptive Biosystem Cytofluor 4000 instrument are as follows:

| Specific Gene Signal: | | Housekeeping Gene Signal: | |
|-----------------------|----------|---------------------------|----------|
| Excitation: | 485/20nm | Excitation: | 560/20nm |
| Emission: | 530/25nm | Emission: | 620/40nm |
| Reads/Well: | 10 | Reads/Well: | 10 |
| Gain: | 40 | Gain: | 45 |
| Temperature: | 25°C | Temperature: | 25°C |

NOTE: Because the optimal gain setting can vary between instruments, adjust the gain as needed to give the best signal/background ratio (sample raw signal divided by the No Target Control signal) or No Target Control sample readings of ~100 RFUs. Fluorescence microplate readers that use a xenon lamp source generally produce higher RFUs. For directly reading the microplates, the probe height of, and how the plate is positioned in, the fluorescence microplate reader may need to be adjusted according to the manufacturer's recommendations.

Stop and Transfer Procedure

1. Prepare 1X Stop Solution (10mM Tris•HCl, pH 8, 10mM EDTA) with RNase-free H₂O. Add 100µl per well with a multichannel pipet.
2. Transfer 100µl of the diluted reactions to a black microplate (*e.g.*, COSTAR (Corning), Cat. No. 3915).
3. Read the microplate using the same parameters as the Direct Read Procedure, but adjust the gain to give No Target Control sample readings of ~100 RFUs (see NOTE above).

In some embodiments, supplementary documentation, such as protocols for ancillary procedures, *e.g.*, for the preparation of additional reagents, or for preparation of

samples for use in the methods of the present invention, are provided. In preferred embodiments, supplementary documentation includes guidelines and lists of precautions provided to facilitate successful use of the methods and kits by unskilled or inexperienced users. In particularly preferred embodiments, supplementary documentation includes a troubleshooting guide, *e.g.*, a guide describing possible problems that may be encountered by users, and providing suggested solutions or corrections to intended to aid the user in resolving or avoiding such problems.

For example, and without intending to limit such supplementary documentation to any particular content, kits of the present invention may comprise any of the following procedures and guidelines:

II. AVOIDANCE OF RNase CONTAMINATION

To avoid RNase contamination during sample preparation and testing, in one embodiment, the user is cautioned to observe the following precautions:

- Wear disposable gloves at all times to avoid contact with samples and reagents.
- Use certified RNase-free disposables, including thin-wall polypropylene tubes and aerosol-barrier pipet tips, for preparing samples and assay reagents, to avoid cross-contamination.
- Use RNase-free (DEPC-treated) H₂O for diluting samples and/or reagents.
- Keep RNA samples and controls on ice during assay setup.

III. SAMPLE AND CONTROL PREPARATION

NOTE: Dilute both standards and samples to concentrations that correspond to a 5- μ l addition per reaction.

Example 1: The concentration of a 5-attomole standard is 1amol/ μ l. 1amol = 10⁻¹⁸ mole = 602,000 molecules.

Example 2: The concentration of a 100-ng sample should be 20ng/ μ l.

A. Control Preparation

No Target Control:

Total RNA Format: tRNA Carrier (20ng/μl)

Cell Lysate Format: 1X Cell Lysis Buffer 1 (dilute 10X Cell Lysis Buffer 1 to 1X
with RNase-free H₂O)

Positive Control: RNA Standard (Std) (100amol/μl *in vitro* transcript)

1. Prepare RNA standards by diluting the positive controls with tRNA Carrier (when running total RNA samples) or with 1X Cell Lysis Buffer 1 [10X Cell Lysis Buffer 1 diluted with RNase-free H₂O] (when running cell lysate samples). The Product Information Sheet included in each kit indicates the recommended standard test levels and preparation methods.
2. Using a fresh set of standards for each run is recommended. Store the standards on ice during reaction setup.

B. Total RNA Sample Preparation

1. Prepare total RNA from cells or tissue according to manufacturer's instructions for the selected preparation method. Recommended methods include TRIZOL (Life Technologies, Rockville, MD), RNEASY (Qiagen, Valencia, CA), and RNA WIZ (Ambion, Austin, TX).
2. Dilute total RNA samples with RNase-free H₂O to the appropriate concentration.

C. Cell Lysate Sample Preparation -- 96-well microplate format

NOTE: This cell lysate detection format is used for adherent cells cultured in 96-well tissue culture microplates. Cells are typically seeded at 10,000–40,000 cells per well. Different seeding densities may be required depending on cell type and/or mRNA expression levels. See Procedural Notes for more details. For cells exhibiting high expression, the following methods can be used to attenuate the signal from the cell lysates:

- plate fewer cells per well;
 - dilute the cell lysates with 1X Cell Lysis Buffer 1 before addition to the reaction (*e.g.*, 2.5µl lysate + 2.5µl 1X Cell Lysis Buffer 1);
 - read the reaction microplate 15–30 minutes after addition of the
- 5 Secondary FRET Reaction Mix instead of the recommended 60–90 minutes;

1. Dilute 10X Cell Lysis Buffer 1 to a 1X concentration with RNase-free H₂O.
2. Using a multichannel pipet, carefully remove the culture medium from the wells of
- 10 adherent cells without disturbing the cell monolayer.
3. Wash the cells once with 200µl PBS (no MgCl₂, no CaCl₂) and carefully remove the residual PBS with the multichannel pipet.
4. Add 40µl 1X Cell Lysis Buffer 1 per well. Lyse cells at room temperature for 3–5 minutes.
- 15 5. Using a multichannel pipet, carefully transfer 25µl of each lysate sample into a 96-well microplate. Avoid transferring cellular material from the bottom of the well.
6. Overlay each lysate sample with 10µl clear CHILLOUT or mineral oil (overlying is not necessary if using a heated-lid thermal cycler).
7. Seal microplate with Thermaseal well tape. Immediately heat lysates at 75–80°C for
- 20 15 minutes in a thermal cycler or oven to inactivate cellular nucleases.
8. During the heating step, proceed with the reaction setup. See DETAILED mRNA INVADER ASSAY PROTOCOL (above) for instructions.
9. After the heat inactivation step, add the lysate samples immediately to the reaction microplate. Alternatively, the lysate samples can be quickly transferred to a –70°C
- 25 freezer for later testing (long-term stability has not been established and may differ for each cell type).

IV. PROCEDURAL NOTES FOR OPERATION OF THE mRNA INVADER ASSAY

1. RNA sample types and optimization of RNA sample amount.

The assay is optimized for performance with total RNA samples prepared from either tissue or cells. Several total RNA preparation methods/kits have been validated for performance in the mRNA INVADER assay:

- TRIZOL (Life Technologies, Rockville, MD)
- RNeasy (Qiagen, Valencia, CA)
- RNA WIZ (Ambion, Austin, TX)

It is important to use a method or kit that minimizes the level of genomic DNA, which can inhibit signal generation. Performance of a preliminary experiment is recommended to determine the amount of total RNA sample (typically 1–200ng, depending on the gene's expression level) that provides the best limit of detection and dynamic range.

The assay has also been validated with lysate samples from a number of cell types. Recommended cell densities in a 96-well tissue culture microplate are 10,000–40,000 cells per well depending on cell type and expression level of the gene of interest. Performance of a preliminary experiment is recommended for any given cell line and/or gene being monitored. Such an experiment should include different cell density levels and/or dilution of the lysate samples with 1X Cell Lysis Buffer 1 (e.g. a 1µl test level is prepared by mixing 1µl lysate sample + 4µl 1X Cell Lysis Buffer 1 for a 5µl sample addition).

2. Dynamic range modulation: variable secondary reaction incubation times.

The length of the secondary reaction incubation time listed in the protocol is sufficient for most analytes. However, the linear detection range (Signal/Background < 15–25) can be adjusted by reading the reaction microplate at variable times after addition of the secondary FRET reagents. For example,

high expression samples can often be detected in 15–30 minutes. The Direct Read method (DETAILED mRNA INVADER ASSAY PROTOCOL, step 7) enables simple optimization of the secondary reaction time as the reaction microplate can be incubated further if an early time read does not provide enough signal from the samples being tested.

Monitoring the secondary reaction fluorescence signal with time can also extend the dynamic range of the assay. The Direct Read method at multiple time points can be applied using low-cost instrumentation. Alternatively, real-time fluorescence instrumentation can be used to achieve comparable dynamic ranges exhibited by other mRNA quantitation methods.

3. Dynamic range modulation: variable sample levels.

While the FRET detection method greatly simplifies the assay, the dynamic range is typically limited to 2–3 logs when using an endpoint read method. However, since mRNA INVADER assay signal is generated linearly with both target level and time, the easiest method for extending the dynamic range beyond 3 logs (as may be required, *e.g.*, for highly induced genes) is to adjust total RNA sample levels. Fold changes in gene expression (treated sample signal divided by untreated sample signal) can be reliably calculated using normalized sample signals. This is accomplished by testing sample levels that give signal within the linear detection range defined by the standard curve. For example, the fold induction for a highly induced sample can be calculated as follows:

Fold induction = (Net Signal for 1 ng treated sample X 100) / Net Signal for 100 ng untreated sample

V. TROUBLESHOOTING GUIDE

| Problem | Possible Solution |
|--|--|
| No signal | <ul style="list-style-type: none"> • Check that the fluorescence microplate reader has been set up correctly and that the appropriate excitation and emission filters are in place. • Perform mRNA INVADER assay with the provided standard as a positive control. • Potential RNase contamination of the samples and reagents. Discard suspect reagents. • Use only reagents and oligonucleotides supplied in the kit. Do not mix reagents or oligonucleotides between kits. |
| High variation between replicates | <ul style="list-style-type: none"> • Always work with master primary and secondary reaction mixes. • Thoroughly mix all master mixes and samples. • Pipet in a similar manner across all the controls and samples. • Calibrate pipets frequently. |
| Lack of low target level detection | <ul style="list-style-type: none"> • Calibrate thermal cycler or heat block. • Minimize assay variability (see above), i.e. CVs are less than 5% for the sample replicates. This is particularly important for detecting low target levels. |
| Lack of discrimination between high signal samples | <ul style="list-style-type: none"> • Decrease secondary reaction incubation time to achieve detection within the linear range of the assay. • Use less total RNA per reaction. • Attenuate cell lysate sample signal (see NOTE, Sample and Control Preparation, Part C). |
| Signal inhibition | <ul style="list-style-type: none"> • Run samples on an agarose gel to check for presence of genomic DNA. Alter the RNA sample isolation method to minimize genomic DNA or presence of other inhibitors. The same isolation procedure should be used throughout an experiment. • If using the cell lysate format, residual PBS can be inhibitory. Be sure to remove residual PBS from the tissue culture microplate. Do not use PBS that contains $MgCl_2$ or $CaCl_2$, which inhibits the assay. |

APPENDIX A:

mRNA INVADER SINGLE ASSAY WORKSHEET

mRNA INVADER ASSAY PROCEDURE

- Prepare samples and controls.
- 5 ▪ Prepare Primary Reaction Mix. Vortex briefly and aliquot 5µl per well.
- Add 5µl sample or control per well and pipet up and down 1–2 times.
- Add 10µl CHILLOUT or mineral oil per well.
- Incubate primary reaction at 60°C for 90 minutes.
- Prepare Secondary FRET Reaction Mix, vortex briefly.
- 10 ▪ Using a multichannel pipet, aliquot 5µl per well and pipet up and down 1–2 times.
- Incubate secondary reaction at 60°C for 60 or 90 minutes.
- Read microplate in fluorescence microplate reader (FAM Dye: Ex. 485 nm/Em. 530 nm).

PRIMARY REACTION MIX

| Reaction Components | 1X Volume | ____X Volume (No. of reactions x 1.25) |
|---|--------------|--|
| RNA Primary Buffer 1 | 4.0µl | |
| ____ Primary Oligos | 0.25µl | |
| T ₁₀ e _{0.1} Buffer | 0.25µl | |
| CLEAVASEenzyme | 0.5µl | |
| Total Mix Volume (1X) | 5.0µl | |

SECONDARY FRET REACTION MIX

| Reaction Components | 1X Volume | ____X Volume (No. of reactions x 1.25) |
|---|--------------|--|
| RNA Secondary Buffer 1 | 2.0µl | |
| ____ Secondary Oligos | 1.5µl | |
| T ₁₀ e _{0.1} Buffer | 1.5µl | |
| Total Mix Volume (1X) | 5.0µl | |

APPENDIX B:

mRNA INVADER BIPLEX ASSAY WORKSHEET

mRNA INVADER ASSAY PROCEDURE

- 5 ▪ Prepare samples and controls.
- Prepare Primary Reaction Mix. Vortex briefly and aliquot 5µl per well.
- Add 5µl sample or control per well and pipet up and down 1–2 times.
- Add 10µl CHILLOUT or mineral oil per well.
- Incubate primary reaction at 60°C for 90 minutes.
- 10 ▪ Prepare Secondary FRET Reaction Mix, vortex briefly.
- Using a multichannel pipet, aliquot 5µl per well and pipet up and down 1–2 times.
- Incubate secondary reaction at 60°C for 60 or 90 minutes.
- Read microplate in fluorescence microplate reader (FAM Dye: Ex. 485 nm/Em. 530 nm and red dye: Ex. 560nm/Em. 620nm).

PRIMARY REACTION MIX

| Reaction Components | 1X Volume | ____X Volume (# reactions x 1.25) |
|----------------------------------|--------------|--------------------------------------|
| RNA Primary Buffer 1 | 4.0µl | |
| ____ Primary Oligos | 0.25µl | |
| ____ Housekeeping Primary Oligos | 0.25µl | |
| CLEAVASE IX enzyme | 0.5µl | |
| Total Mix Volume (1X) | 5.0µl | |

SECONDARY FRET REACTION MIX

| Reaction Components | 1X Volume | ____X Volume (# reactions x 1.25) |
|------------------------------------|--------------|--------------------------------------|
| RNA Secondary Buffer 1 | 2.0µl | |
| ____ Secondary Oligos | 1.5µl | |
| ____ Housekeeping Secondary Oligos | 1.5µl | |
| Total Mix Volume (1X) | 5.0µl | |

TABLE 2

| <u>Klenow</u> | <u>Kcat (s⁻¹)</u> | <u>Km(dNTP)</u> (μ M) | <u>Kd (nM)</u> | <u>Relative DNA</u> <u>affinity</u> | <u>Reference</u> | <u>Taq Pol</u> |
|---------------|------------------------------|-------------------------------|----------------|--|------------------|----------------|
| Wild-Type | 2.4 | 2.8 | 8 | 1 | 2 | Wild-Type |
| S610A | n. d. | - | n. d. | - | 5 | S515 |
| R668A | 0.006 | 6.5 | 140, 150 | 0.06, 0.05 | 1,2 | R573* |
| N678A | n. d. | - | n. d. | - | 5 | N583 |
| E710A | 0.1 | 15 | 250 | 0.03 | 2 | E615* |
| E710D | 1.7 | 7.7 | 110 | 0.07 | 2 | E615* |
| K758A | 0.131 | 15.6 | - | 0.63 | 4 | K663 |
| K758R | 2.0 | 2.1 | - | 1.125 | 4 | K663 |
| Y766S | 0.8 | 6.4 | 13 | 0.4, 0.6 | 1,2 | Y671 |
| R841A | 0.3 | 9.8 | 40, 53 | 0.2 | 1,2 | R746* |
| N845A | 1.0 | 23 | 8, 5 | 1.0, 1.7 | 1,2 | N750* |
| N845Q | 0.03 | 1.7 | 80, 55 | 0.1, 0.2 | 1,2 | N750* |
| Q849A | 0.02 | 3.8 | 100, 160 | 0.08, 0.05 | 1,2 | Q754 |
| Q849E | 0.001 | n. d. | 90, 91 | 0.09 | 1,2 | Q754 |
| H881A | 0.3 | 3.3 | 20, 28 | 0.4, 0.3 | 1,2 | H784* |
| D882N | <0.0001 | n. d. | 30 | 0.6 | 2 | D785 |
| D882S | 0.001 | 7.5 | 0.9 | 9 | 2 | D785 |

References:

1. JBC (1990) 265:14579-14591
2. JBC (1992) 267:8417-8428
3. Eur. J. Biochem (1993) 214:59-65
4. JBC (1994) 269:13259-13265
5. Nature (1996) 382:278-281

TABLE 3: Rational mutations in the polymerase region

A. DNA activity table

| | IdT | %Tth | %Taq4M | HP | X |
|-----------------------|--------------|-------------|------------|-------------|--------------|
| <u>Tth DN RX HT</u> | <u>31.91</u> | <u>100%</u> | <u>83%</u> | <u>3.81</u> | <u>101.9</u> |
| Tth DN RX HT H641A | 23.61 | 74% | 62% | 5.32 | 221.24 |
| Tth DN RX HT R748A | 22.1 | 69% | 58% | 4.39 | 88.17 |
| Tth DN RX HT H786A | 34.31 | 108% | 90% | 7.75 | 185.35 |
| Tth DN RX HT | 32.1 | 101% | 84% | 5.7 | 332.8 |

| | | | | | |
|---|-------|------|------|--------|---------|
| H786A/G506K/ Q509K (AKK) | | | | | |
| Taq DN RX HT W417L/G418K/ E507Q/H784A (Taq 4M) | 38.23 | 120% | 100% | 68.21 | 1100.18 |
| Taq 4M G504K | 36.04 | 113% | 94% | 31.76 | 417.40 |
| Taq 4M H639A | 42.95 | 135% | 112% | 91.46 | 2249.67 |
| Taq 4M R587A | 44.78 | 140% | 117% | 143.0 | 252.69 |
| Taq DN RX HT W417L/G418K/ G499R/A502K/ I503L/K504N/ E07K/H784A (Taq8M) | 43.95 | 138% | 115% | 122.53 | 346.56 |
| TaqSS R677A | 32.3 | 101% | 84% | 206.9 | 2450.0 |

B. RNA activity table

| | IrT1 | %Tth | %Taq4M |
|---|-------------|-------------|------------|
| <u>Tth DN RX HT</u> | <u>0.89</u> | <u>100%</u> | <u>34%</u> |
| Tth DN RX HT H641A | 1.18 | 133% | 45% |
| Tth DN RX HT R748A | 1.34 | 151% | 51% |
| Tth DN RX HT H786A | 1.31 | 147% | 49% |
| Tth DN RX HT H786A/G506K/ Q509K (AKK) | 1.59 | 179% | 60% |
| Taq DN RX HT W417L/G418K/ E507Q/H784A (Taq 4M) | 2.65 | 298% | 100% |
| Taq 4M G504K | 2.76 | 310% | 114% |
| Taq 4M H639A | 3.89 | 437% | 147% |
| Taq 4M R587A | 3.13 | 352% | 118% |
| Taq DN RX HT W417L/G418K/ G499R/A502K/ I503L/K504N/ E07K/H784A (Taq8M) | 4.00 | 450% | 151% |
| TaqSS R677A | 2.22 | 249% | 84% |

TABLE 4: Rational arch mutations**DNA activity table**

| | IdT | %Tth | %Taq4M | HP | X |
|-----------------------|-------|------|--------|-------|--------|
| Taq 4M P88E/P90E | 10.20 | 32% | 27% | 2.00 | 97.00 |
| Taq 4M G80E | 26.30 | 82% | 69% | 103.6 | 2900 |
| Taq 4M L109F/A110T | 36.45 | 114% | 95% | 19.71 | 749.69 |

5

RNA activity table

| | IrT1 | %Tth | %Taq4M | |
|-----------------------|-----------------------|------|--------|------|
| Taq 4M P88E/P90E | Taq 4M P88E/P90E | 0.10 | 11% | 4% |
| Taq 4M G80E | Taq 4M G80E | 3.11 | 349% | 117% |
| Taq 4M L109F/A110T | Taq 4M L109F/A110T | 2.45 | 275% | 92% |

10

TABLE 5: Arch/thumb combinations**DNA activity table**

| | IdT | %Tth | %Taq4M | HP | X |
|--|-------|------|--------|--------|--------|
| Taq W417L/ G418K/E507K/ H784A/L109F/ A110T/G499R/ A502K/I503L/ G504K/E507K/ T514S (Taq SS) | 63.33 | 198% | 166% | 177.05 | 202.32 |
| Taq P88E/P90E/ W417L/G418K/ G499R/A502K/ I503L/G504K/ E507K/T514S/ H784A | 36.48 | 114% | 95% | 9.44 | 70.35 |

15

RNA activity table

| | IrT1 | %Tth | %Taq4M |
|--|------|------|--------|
| Taq W417L/ G418K/E507K/ H784A/L109F/ A110T/G499R/ A502K/I503L/ G504K/E507K/ T514S (Taq SS) | 3.16 | 355% | 119% |
| Taq P88E/P90E/ W417L/G418K/ G499R/A502K/ I503L/G504K/ E507K/T514S/ H784A | 0.22 | 25% | 8% |

TABLE 6: Helix-hairpin-helix random mutagenesis

DNA activity table

| | IdT | %Tth | %Taq4M | HP | X |
|------------------------------------|------|------|--------|------|--------|
| TaqSS K198N | 23.4 | 73% | 61% | 25.7 | 1233.1 |
| TaqSS A205Q | 25.6 | 80% | 67% | 13.4 | 699.1 |
| TaqSS T204P | 11.2 | 35% | 29% | 1.9 | 209.4 |
| TaqSS I200M/A205G | 16.8 | 53% | 44% | 7.8 | 597.2 |
| TaqSS K203N | 25.9 | 81% | 68% | 36.6 | 1429.8 |
| Tth DN RX HT H786A/P197R/K200R | 10.7 | 33% | 28% | 3.2 | 66.3 |
| Tth DN RX HT H786A/K205Y | 11.5 | 36% | 30% | 6.1 | 327.5 |
| Tth DN RX HT H786A/G203R | 18.3 | 57% | 48% | 2.1 | 98.8 |

RNA activity table

| | IrT1 | %Tth | %Taq4M |
|------------------------------------|------|------|--------|
| TaqSS K198N | 1.22 | 137% | 46% |
| TaqSS A205Q | 0.62 | 70% | 23% |
| TaqSS T204P | 0.36 | 40% | 14% |
| TaqSS I200M/A205G | 0.77 | 87% | 29% |
| TaqSS K203N | 2.09 | 235% | 79% |
| Tth DN RX HT H786A/P197R/K200R | 0.47 | 52% | 18% |
| Tth DN RX HT H786A/K205Y | 0.68 | 77% | 26% |
| Tth DN RX HT H786A/G203R | 1.61 | 180% | 61% |

5 **TABLE 7: Random thumb mutations**
DNA activity table

| | IdT | %Tth | %Taq4M | HP | X |
|---|-------|------|--------|--------|---------|
| Taq DN RX HT W417L/G418K/ E507K/H784A/G499R/ A502K/K504N/ (M1-13) | 59.96 | 188% | 157% | 133.65 | 907.41 |
| Taq DN RX HT W417L/G418K/ /H784A/L500I/Q507H A502K/G504K (M1-36) | 46.74 | 146% | 122% | 123.11 | 822.61 |
| Taq DN RX HT W417L/G418K/G499R/ A502K/G504K/E507K/ H784A/T514S (M2-24) | 85.7 | 269% | 224% | 369.96 | 3752.12 |
| Taq DN RX HT W417L/G418K/G499R/ A502K/G504K/E507K/ H784A/V518L (M2-06) | 76.7 | 240% | 201% | 355.87 | 2038.19 |

RNA activity table

| | IrT1 | %Tth | %Taq4M |
|---|------|------|--------|
| Taq DN RX HT W417L/G418K/ E507K/H784A/G499R/ A502K/K504N/ (M1-13) | 2.55 | 287% | 96% |
| Taq DN RX HT W417L/G418K/ /H784A/L500I/Q507H A502K/G504K (M1-36) | 2.71 | 304% | 102% |
| Taq DN RX HT W417L/G418K/G499R/ A502K/G504K/E507K/ H784A/T514S (M2-24) | 4.43 | 498% | 167% |
| Taq DN RX HT W417L/G418K/G499R/ A502K/G504K/E507K/ H784A/V518L (M2-06) | 3.56 | 400% | 134% |

5 TABLE 8: Chimeric mutants

A. DNA activity table

| | IdT2 | %TthAKK | HP | X |
|--------------------|--------------|-------------|----------|------------|
| <u>TthAKK</u> | <u>34.18</u> | <u>100%</u> | <u>5</u> | <u>393</u> |
| Taq 4M G504K | 40.19 | 105% | 28 | 1991 |
| Tfi DN 2M | 36.60 | 106% | 289 | 1326 |
| Tsc DN 2M | 25.49 | 75% | 283 | 2573 |
| TaqTthAKK | 63.89 | 187% | 32 | 1658 |
| TthTaq 4M G504K | 25.03 | 73% | 8 | 627 |
| TfiTthAKK | 34.13 | 100% | 15 | 459 |
| TscTthAKK | 35.23 | 103% | 29 | 2703 |
| TfiTaq 4M G504K | 35.69 | 104% | 37 | 872 |
| TscTaq 4M G504K | 30.04 | 88% | 25 | 2008 |

B. RNA activity table

| | IrT3 | %TthAKK |
|---------------|-------------|-------------|
| <u>TthAKK</u> | <u>2.27</u> | <u>100%</u> |
| Taq 4M G504K | 2.31 | 102% |
| Tfi DN 2M | 0.20 | 9% |
| Tsc DN 2M | 0.29 | 13% |

| | | |
|--------------------|------|-------|
| TaqTthAKK | 6.81 | 300% |
| TthTaq 4M G504K | 1.09 | 48% |
| TfiTthAKK | 1.24 | 55% |
| TscTthAKK | 9.65 | 4.25% |
| TfiTaq 4M G504K | 1.05 | 46% |
| TscTaq 4M G504K | 2.95 | 130% |

All publications and patents mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in the relevant fields are intended to be within the scope of the following claims.

SEQUENCE LISTING

<110> Third Wave Technologies

<120> Detection of RNA

<130> FORS

<160> 708

<170> PatentIn version 3.0

<210> 1

<211> 834

<212> PRT

<213> Thermus aquaticus

<400> 1

```

Met Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu
1          5          10          15

Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu Lys Gly
          20          25          30

Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala
          35          40          45

Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala Val Phe
          50          55          60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu
65          70          75          80

Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
          85          90          95

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr Arg Leu
          100          105          110

Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys
          115          120          125

Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Arg
          130          135          140

Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His Pro Glu
145          150          155          160

Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly Leu Arg
          165          170          175

Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro Ser Asp
          180          185          190

Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu Lys Leu
          195          200          205

```


Lys Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His
 515 520 525
 Pro Ile Val Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys Leu Lys
 530 535 540
 Asn Thr Tyr Val Asp Pro Leu Pro Ser Leu Val His Pro Arg Thr Gly
 545 550 555 560
 Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu
 565 570 575
 Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu
 580 585 590
 Gly Gln Arg Ile Arg Arg Ala Phe Val Ala Glu Ala Gly Trp Ala Leu
 595 600 605
 Val Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu
 610 615 620
 Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys Asp Ile
 625 630 635 640
 His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu Ala Val
 645 650 655
 Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly Val Leu
 660 665 670
 Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr
 675 680 685
 Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys
 690 695 700
 Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys Arg Gly
 705 710 715 720
 Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn
 725 730 735
 Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn
 740 745 750
 Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val
 755 760 765
 Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu Leu Gln
 770 775 780
 Val His Asp Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala Glu Glu
 785 790 795 800
 Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro Leu Ala
 805 810 815

Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala
820 825 830

Lys Gly

<210> 2
<211> 839
<212> PRT
<213> Artificial

<220>
<223> Synthetic

<400> 2

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1 5 10 15

Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
20 25 30

Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
35 40 45

Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
50 55 60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
65 70 75 80

Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
85 90 95

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
100 105 110

Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
115 120 125

Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
130 135 140

Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
145 150 155 160

Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
165 170 175

Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
180 185 190

Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
195 200 205

Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg

| | | |
|--|-----|-----|
| 515 | 520 | 525 |
| Ile Val Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys Leu Lys Asn 530 535 540 | | |
| Thr Tyr Val Asp Pro Leu Pro Ser Leu Val His Pro Arg Thr Gly Arg 545 550 555 560 | | |
| Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser 565 570 575 | | |
| Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly 580 585 590 | | |
| Gln Arg Ile Arg Arg Ala Phe Val Ala Glu Ala Gly Trp Ala Leu Val 595 600 605 | | |
| Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser 610 615 620 | | |
| Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys Asp Ile His 625 630 635 640 | | |
| Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu Ala Val Asp 645 650 655 | | |
| Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly Val Leu Tyr 660 665 670 | | |
| Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu 675 680 685 | | |
| Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val 690 695 700 | | |
| Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys Arg Gly Tyr 705 710 715 720 | | |
| Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn Ala 725 730 735 | | |
| Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met 740 745 750 | | |
| Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys 755 760 765 | | |
| Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu Leu Gln Val 770 775 780 | | |
| His Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala Glu Glu Val 785 790 795 800 | | |
| Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro Leu Ala Val 805 810 815 | | |
| Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala Lys | | |

| | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Lys | Leu | Ser | Trp | Asp | Leu | Ala | Lys | Val | Arg | Thr | Asp | Leu | Pro | Leu | Glu | 245 | 250 | 255 |
| Val | Asp | Phe | Ala | Lys | Arg | Arg | Glu | Pro | Asp | Arg | Glu | Arg | Leu | Arg | Ala | 260 | 265 | 270 |
| Phe | Leu | Glu | Arg | Leu | Glu | Phe | Gly | Ser | Leu | Leu | His | Glu | Phe | Gly | Leu | 275 | 280 | 285 |
| Leu | Glu | Ser | Pro | Lys | Ala | Leu | Glu | Glu | Ala | Pro | Trp | Pro | Pro | Pro | Glu | 290 | 295 | 300 |
| Gly | Ala | Phe | Val | Gly | Phe | Val | Leu | Ser | Arg | Lys | Glu | Pro | Met | Trp | Ala | 305 | 310 | 315 |
| Asp | Leu | Leu | Ala | Leu | Ala | Ala | Ala | Arg | Gly | Gly | Arg | Val | His | Arg | Ala | 325 | 330 | 335 |
| Ala | Asp | Pro | Leu | Ala | Gly | Leu | Lys | Asp | Leu | Lys | Glu | Val | Arg | Gly | Leu | 340 | 345 | 350 |
| Leu | Ala | Lys | Asp | Leu | Ala | Val | Leu | Ala | Ser | Arg | Glu | Gly | Leu | Asp | Leu | 355 | 360 | 365 |
| Val | Pro | Gly | Asp | Asp | Pro | Met | Leu | Leu | Ala | Tyr | Leu | Leu | Asp | Pro | Ser | 370 | 375 | 380 |
| Asn | Thr | Thr | Pro | Glu | Gly | Val | Ala | Arg | Arg | Tyr | Gly | Gly | Glu | Trp | Thr | 385 | 390 | 395 |
| Glu | Asp | Ala | Ala | His | Arg | Ala | Leu | Leu | Ser | Glu | Arg | Leu | His | Arg | Asn | 405 | 410 | 415 |
| Leu | Leu | Lys | Arg | Leu | Glu | Gly | Glu | Glu | Lys | Leu | Leu | Trp | Leu | Tyr | His | 420 | 425 | 430 |
| Glu | Val | Glu | Lys | Pro | Leu | Ser | Arg | Val | Leu | Ala | His | Met | Glu | Ala | Thr | 435 | 440 | 445 |
| Gly | Val | Arg | Arg | Asp | Val | Ala | Tyr | Leu | Gln | Ala | Leu | Ser | Leu | Glu | Leu | 450 | 455 | 460 |
| Ala | Glu | Glu | Ile | Arg | Arg | Leu | Glu | Glu | Glu | Val | Phe | Arg | Leu | Ala | Gly | 465 | 470 | 475 |
| His | Pro | Phe | Asn | Leu | Asn | Ser | Arg | Asp | Gln | Leu | Glu | Arg | Val | Leu | Phe | 485 | 490 | 495 |
| Asp | Glu | Leu | Arg | Leu | Pro | Ala | Leu | Gly | Lys | Thr | Gln | Lys | Thr | Gly | Lys | 500 | 505 | 510 |
| Arg | Ser | Thr | Ser | Ala | Ala | Val | Leu | Glu | Ala | Leu | Arg | Glu | Ala | His | Pro | 515 | 520 | 525 |
| Ile | Val | Glu | Lys | Ile | Leu | Gln | His | Arg | Glu | Leu | Thr | Lys | Leu | Lys | Asn | 530 | 535 | 540 |

<210> 5
 <211> 839
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 5

```

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1          5          10          15

Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
          20          25          30

Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
          35          40          45

Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
          50          55          60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
65          70          75          80

Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
          85          90          95

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
          100         105         110

Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
          115         120         125

Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
          130         135         140

Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
145          150         155         160

Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
          165         170         175

Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
          180         185         190

Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
          195         200         205

Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
          210         215         220

Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
225          230         235         240

Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu

```


<211> 839
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 6

```

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1          5          10          15

Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
          20          25          30

Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
          35          40          45

Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
          50          55          60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
65          70          75          80

Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
          85          90          95

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
          100          105          110

Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
          115          120          125

Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
          130          135          140

Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
145          150          155          160

Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
          165          170          175

Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
          180          185          190

Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
          195          200          205

Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
          210          215          220

Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
225          230          235          240

Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
          245          250          255

```

09864633-02401

<213> Artificial

<220>

<223> Synthetic

<400> 7

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1 5 10 15
Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
20 25 30
Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
35 40 45
Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
50 55 60
Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
65 70 75 80
Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
85 90 95
Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
100 105 110
Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
115 120 125
Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
130 135 140
Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
145 150 155 160
Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
165 170 175
Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
180 185 190
Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
195 200 205
Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
210 215 220
Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
225 230 235 240
Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
245 250 255
Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
260 265 270

<220>

<223> Synthetic

<400> 8

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1 5 10 15
Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
20 25 30
Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
35 40 45
Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
50 55 60
Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
65 70 75 80
Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
85 90 95
Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
100 105 110
Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
115 120 125
Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
130 135 140
Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
145 150 155 160
Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
165 170 175
Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
180 185 190
Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
195 200 205
Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
210 215 220
Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
225 230 235 240
Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
245 250 255
Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
260 265 270
Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu

| 580 | 585 | 590 |
|--|-----|-----|
| Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val 595 | 600 | 605 |
| Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser 610 | 615 | 620 |
| Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His 625 | 630 | 635 |
| Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp 645 | 650 | 655 |
| Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr 660 | 665 | 670 |
| Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu 675 | 680 | 685 |
| Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val 690 | 695 | 700 |
| Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr 705 | 710 | 715 |
| Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala 725 | 730 | 735 |
| Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met 740 | 745 | 750 |
| Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys 755 | 760 | 765 |
| Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val 770 | 775 | 780 |
| His Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val 785 | 790 | 795 |
| Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val 805 | 810 | 815 |
| Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys 820 | 825 | 830 |
| Glu His His His His His His 835 | | |
| <210> 9 | | |
| <211> 839 | | |
| <212> PRT | | |
| <213> Artificial | | |
| <220> | | |
| <223> Synthetic | | |

Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu
 290 295 300
 Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala
 305 310 315 320
 Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala
 325 330 335
 Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val Arg Gly Leu
 340 345 350
 Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly Leu Asp Leu
 355 360 365
 Val Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser
 370 375 380
 Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr
 385 390 395 400
 Glu Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn
 405 410 415
 Leu Trp Gly Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg
 420 425 430
 Glu Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr
 435 440 445
 Gly Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val
 450 455 460
 Ala Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly
 465 470 475 480
 His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe
 485 490 495
 Asp Glu Leu Gly Leu Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys
 500 505 510
 Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro
 515 520 525
 Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser
 530 535 540
 Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg
 545 550 555 560
 Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser
 565 570 575
 Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly
 580 585 590

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Asn | Ser | Glu | Ala | Met | Leu | Pro | Leu | Phe | Glu | Pro | Lys | Gly | Arg | Val | 1 | 5 | 10 | 15 |
| Leu | Leu | Val | Asp | Gly | His | His | Leu | Ala | Tyr | Arg | Thr | Phe | Phe | Ala | Leu | 20 | 25 | 30 | |
| Lys | Gly | Leu | Thr | Thr | Ser | Arg | Gly | Glu | Pro | Val | Gln | Ala | Val | Tyr | Gly | 35 | 40 | 45 | |
| Phe | Ala | Lys | Ser | Leu | Leu | Lys | Ala | Leu | Lys | Glu | Asp | Gly | Tyr | Lys | Ala | 50 | 55 | 60 | |
| Val | Phe | Val | Val | Phe | Asp | Ala | Lys | Ala | Pro | Ser | Phe | Arg | His | Glu | Ala | 65 | 70 | 75 | 80 |
| Tyr | Glu | Ala | Tyr | Lys | Ala | Gly | Arg | Ala | Pro | Thr | Pro | Glu | Asp | Phe | Pro | 85 | 90 | 95 | |
| Arg | Gln | Leu | Ala | Leu | Ile | Lys | Glu | Leu | Val | Asp | Leu | Leu | Gly | Phe | Thr | 100 | 105 | 110 | |
| Arg | Leu | Glu | Val | Pro | Gly | Tyr | Glu | Ala | Asp | Asp | Val | Leu | Ala | Thr | Leu | 115 | 120 | 125 | |
| Ala | Lys | Lys | Ala | Glu | Lys | Glu | Gly | Tyr | Glu | Val | Arg | Ile | Leu | Thr | Ala | 130 | 135 | 140 | |
| Asp | Arg | Asp | Leu | Tyr | Gln | Leu | Val | Ser | Asp | Arg | Val | Ala | Val | Leu | His | 145 | 150 | 155 | 160 |
| Pro | Glu | Gly | His | Leu | Ile | Thr | Pro | Glu | Trp | Leu | Trp | Glu | Lys | Tyr | Gly | 165 | 170 | 175 | |
| Leu | Arg | Pro | Glu | Gln | Trp | Val | Asp | Phe | Arg | Ala | Leu | Val | Gly | Asp | Pro | 180 | 185 | 190 | |
| Ser | Asp | Asn | Leu | Pro | Gly | Val | Lys | Gly | Ile | Gly | Glu | Lys | Thr | Ala | Leu | 195 | 200 | 205 | |
| Lys | Leu | Leu | Lys | Glu | Trp | Gly | Ser | Leu | Glu | Asn | Leu | Leu | Lys | Asn | Leu | 210 | 215 | 220 | |
| Asp | Arg | Val | Lys | Pro | Glu | Asn | Val | Arg | Glu | Lys | Ile | Lys | Ala | His | Leu | 225 | 230 | 235 | 240 |
| Glu | Asp | Leu | Arg | Leu | Ser | Leu | Glu | Leu | Ser | Arg | Val | Arg | Thr | Asp | Leu | 245 | 250 | 255 | |
| Pro | Leu | Glu | Val | Asp | Leu | Ala | Gln | Gly | Arg | Glu | Pro | Asp | Arg | Glu | Gly | 260 | 265 | 270 | |
| Leu | Arg | Ala | Phe | Leu | Glu | Arg | Leu | Glu | Phe | Gly | Ser | Leu | Leu | His | Glu | 275 | 280 | 285 | |
| Phe | Gly | Leu | Leu | Glu | Ala | Pro | Ala | Pro | Leu | Glu | Glu | Ala | Pro | Trp | Pro | 290 | 295 | 300 | |

Leu Leu Val Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala
 610 615 620
 His Leu Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg
 625 630 635 640
 Asp Ile His Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu
 645 650 655
 Ala Val Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly
 660 665 670
 Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile
 675 680 685
 Pro Tyr Glu Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe
 690 695 700
 Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg
 705 710 715 720
 Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp
 725 730 735
 Leu Glu Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala
 740 745 750
 Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala
 755 760 765
 Met Val Lys Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu
 770 775 780
 Leu Gln Val His Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala
 785 790 795 800
 Glu Ala Val Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro
 805 810 815
 Leu Ala Val Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu
 820 825 830
 Ser Ala Lys Glu His His His His His His
 835 840
 <210> 11
 <211> 842
 <212> PRT
 <213> Artificial
 <220>
 <223> Synthetic
 <400> 11
 Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val

| | | | |
|---|-----|-----|-----|
| 1 | 5 | 10 | 15 |
| Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu | 20 | 25 | 30 |
| Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly | 35 | 40 | 45 |
| Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala | 50 | 55 | 60 |
| Val Phe Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala | 65 | 70 | 75 |
| Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro | 85 | 90 | 95 |
| Arg Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr | 100 | 105 | 110 |
| Arg Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu | 115 | 120 | 125 |
| Ala Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala | 130 | 135 | 140 |
| Asp Arg Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His | 145 | 150 | 155 |
| Pro Glu Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly | 165 | 170 | 175 |
| Leu Arg Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro | 180 | 185 | 190 |
| Ser Asp Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu | 195 | 200 | 205 |
| Lys Leu Leu Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu | 210 | 215 | 220 |
| Asp Arg Val Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu | 225 | 230 | 235 |
| Glu Asp Leu Arg Leu Ser Leu Glu Leu Ser Arg Val Arg Thr Asp Leu | 245 | 250 | 255 |
| Pro Leu Glu Val Asp Leu Ala Gln Gly Arg Glu Pro Asp Arg Glu Gly | 260 | 265 | 270 |
| Leu Arg Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu | 275 | 280 | 285 |
| Phe Gly Leu Leu Glu Ala Pro Ala Pro Leu Glu Glu Ala Pro Trp Pro | 290 | 295 | 300 |
| Pro Pro Glu Gly Ala Phe Val Gly Phe Val Leu Ser Arg Pro Glu Pro | | | |

| | | |
|---|-----|-------------|
| 610 | 615 | 620 |
| His Leu Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys | | |
| 625 | 630 | 635 640 |
| Asp Ile His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu | | |
| | 645 | 650 655 |
| Ala Val Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly | | |
| | 660 | 665 670 |
| Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile | | |
| | 675 | 680 685 |
| Pro Tyr Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe | | |
| | 690 | 695 700 |
| Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys | | |
| | 705 | 710 715 720 |
| Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp | | |
| | 725 | 730 735 |
| Leu Asn Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala | | |
| | 740 | 745 750 |
| Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala | | |
| | 755 | 760 765 |
| Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu | | |
| | 770 | 775 780 |
| Leu Gln Val His Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala | | |
| | 785 | 790 795 800 |
| Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro | | |
| | 805 | 810 815 |
| Leu Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu | | |
| | 820 | 825 830 |
| Ser Ala Lys Gly His His His His His His | | |
| | 835 | 840 |
| <210> 12 | | |
| <211> 833 | | |
| <212> PRT | | |
| <213> Artificial | | |
| <220> | | |
| <223> Synthetic | | |
| <400> 12 | | |
| Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu | | |
| 1 | 5 | 10 15 |

Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
 20 25 30
 Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
 35 40 45
 Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
 50 55 60
 Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
 65 70 75 80
 Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
 85 90 95
 Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
 100 105 110
 Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
 115 120 125
 Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
 130 135 140
 Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
 145 150 155 160
 Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
 165 170 175
 Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
 180 185 190
 Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
 195 200 205
 Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
 210 215 220
 Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
 225 230 235 240
 Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
 245 250 255
 Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
 260 265 270
 Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
 275 280 285
 Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu
 290 295 300
 Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala
 305 310 315 320

| | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asp | Leu | Leu | Ala | Leu | Ala | Ala | Ala | Arg | Gly | Gly | Arg | Val | His | Arg | Ala | 325 | 330 | 335 |
| Ala | Asp | Pro | Leu | Ala | Gly | Leu | Lys | Asp | Leu | Lys | Glu | Val | Arg | Gly | Leu | 340 | 345 | 350 |
| Leu | Ala | Lys | Asp | Leu | Ala | Val | Leu | Ala | Ser | Arg | Glu | Gly | Leu | Asp | Leu | 355 | 360 | 365 |
| Val | Pro | Gly | Asp | Asp | Pro | Met | Leu | Leu | Ala | Tyr | Leu | Leu | Asp | Pro | Ser | 370 | 375 | 380 |
| Asn | Thr | Thr | Pro | Glu | Gly | Val | Ala | Arg | Arg | Tyr | Gly | Gly | Glu | Trp | Thr | 385 | 390 | 395 |
| Glu | Asp | Ala | Ala | His | Arg | Ala | Leu | Leu | Ser | Glu | Arg | Leu | His | Arg | Asn | 405 | 410 | 415 |
| Leu | Leu | Lys | Arg | Leu | Glu | Gly | Glu | Glu | Lys | Leu | Leu | Trp | Leu | Tyr | His | 420 | 425 | 430 |
| Glu | Val | Glu | Lys | Pro | Leu | Ser | Arg | Val | Leu | Ala | His | Met | Glu | Ala | Thr | 435 | 440 | 445 |
| Gly | Val | Arg | Arg | Asp | Val | Ala | Tyr | Leu | Gln | Ala | Leu | Ser | Leu | Glu | Leu | 450 | 455 | 460 |
| Ala | Glu | Glu | Ile | Arg | Arg | Leu | Glu | Glu | Glu | Val | Phe | Arg | Leu | Ala | Gly | 465 | 470 | 475 |
| His | Pro | Phe | Asn | Leu | Asn | Ser | Arg | Asp | Gln | Leu | Glu | Arg | Val | Leu | Phe | 485 | 490 | 495 |
| Asp | Glu | Leu | Arg | Leu | Pro | Ala | Leu | Gly | Lys | Thr | Gln | Lys | Thr | Gly | Lys | 500 | 505 | 510 |
| Arg | Ser | Thr | Ser | Ala | Ala | Val | Leu | Glu | Ala | Leu | Arg | Glu | Ala | His | Pro | 515 | 520 | 525 |
| Ile | Val | Glu | Lys | Ile | Leu | Gln | His | Arg | Glu | Leu | Thr | Lys | Leu | Lys | Asn | 530 | 535 | 540 |
| Thr | Tyr | Val | Asp | Pro | Leu | Pro | Ser | Leu | Val | His | Pro | Arg | Thr | Gly | Arg | 545 | 550 | 555 |
| Leu | His | Thr | Arg | Phe | Asn | Gln | Thr | Ala | Thr | Ala | Thr | Gly | Arg | Leu | Ser | 565 | 570 | 575 |
| Ser | Ser | Asp | Pro | Asn | Leu | Gln | Asn | Ile | Pro | Val | Arg | Thr | Pro | Leu | Gly | 580 | 585 | 590 |
| Gln | Arg | Ile | Arg | Arg | Ala | Phe | Val | Ala | Glu | Ala | Gly | Trp | Ala | Leu | Val | 595 | 600 | 605 |
| Ala | Leu | Asp | Tyr | Ser | Gln | Ile | Glu | Leu | Arg | Val | Leu | Ala | His | Leu | Ser | 610 | 615 | 620 |

Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys Asp Ile His
625 630 635 640

Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu Ala Val Asp
645 650 655

Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly Val Leu Tyr
660 665 670

Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
675 680 685

Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
690 695 700

Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys Arg Gly Tyr
705 710 715 720

Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn Ala
725 730 735

Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
740 745 750

Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
755 760 765

Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu Leu Gln Val
770 775 780

His Asp Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala Glu Glu Val
785 790 795 800

Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Gly

<210> 13
<211> 833
<212> PRT
<213> Artificial

<220>
<223> Synthetic

<400> 13

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1 5 10 15

Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
20 25 30

| | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Gly | Leu | Thr | Thr | Ser | Arg | Gly | Glu | Pro | Val | Gln | Ala | Val | Tyr | Gly | Phe | 35 | 40 | 45 |
| Ala | Lys | Ser | Leu | Leu | Lys | Ala | Leu | Lys | Glu | Asp | Gly | Asp | Ala | Val | Ile | 50 | 55 | 60 |
| Val | Val | Phe | Asp | Ala | Lys | Ala | Pro | Ser | Phe | Arg | His | Glu | Ala | Tyr | Gly | 65 | 70 | 75 |
| Gly | Tyr | Lys | Ala | Gly | Arg | Ala | Pro | Thr | Pro | Glu | Asp | Phe | Pro | Arg | Gln | 85 | 90 | 95 |
| Leu | Ala | Leu | Ile | Lys | Glu | Leu | Val | Asp | Leu | Leu | Gly | Leu | Ala | Arg | Leu | 100 | 105 | 110 |
| Glu | Val | Pro | Gly | Tyr | Glu | Ala | Asp | Asp | Val | Leu | Ala | Ser | Leu | Ala | Lys | 115 | 120 | 125 |
| Lys | Ala | Glu | Lys | Glu | Gly | Tyr | Glu | Val | Arg | Ile | Leu | Thr | Ala | Asp | Lys | 130 | 135 | 140 |
| Asp | Leu | Tyr | Gln | Leu | Leu | Ser | Asp | Arg | Ile | His | Val | Leu | His | Pro | Glu | 145 | 150 | 155 |
| Gly | Tyr | Leu | Ile | Thr | Pro | Ala | Trp | Leu | Trp | Glu | Lys | Tyr | Gly | Leu | Arg | 165 | 170 | 175 |
| Pro | Asp | Gln | Trp | Ala | Asp | Tyr | Arg | Ala | Leu | Thr | Gly | Asp | Glu | Ser | Asp | 180 | 185 | 190 |
| Asn | Leu | Pro | Gly | Val | Lys | Gly | Ile | Gly | Glu | Lys | Thr | Ala | Arg | Lys | Leu | 195 | 200 | 205 |
| Leu | Glu | Glu | Trp | Gly | Ser | Leu | Glu | Ala | Leu | Leu | Lys | Asn | Leu | Asp | Arg | 210 | 215 | 220 |
| Leu | Lys | Pro | Ala | Ile | Arg | Glu | Lys | Ile | Leu | Ala | His | Met | Asp | Asp | Leu | 225 | 230 | 235 |
| Lys | Leu | Ser | Trp | Asp | Leu | Ala | Lys | Val | Arg | Thr | Asp | Leu | Pro | Leu | Glu | 245 | 250 | 255 |
| Val | Asp | Phe | Ala | Lys | Arg | Arg | Glu | Pro | Asp | Arg | Glu | Arg | Leu | Arg | Ala | 260 | 265 | 270 |
| Phe | Leu | Glu | Arg | Leu | Glu | Phe | Gly | Ser | Leu | Leu | His | Glu | Phe | Gly | Leu | 275 | 280 | 285 |
| Leu | Glu | Ser | Pro | Lys | Ala | Leu | Glu | Glu | Ala | Pro | Trp | Pro | Pro | Pro | Glu | 290 | 295 | 300 |
| Gly | Ala | Phe | Val | Gly | Phe | Val | Leu | Ser | Arg | Lys | Glu | Pro | Met | Trp | Ala | 305 | 310 | 315 |
| Asp | Leu | Leu | Ala | Leu | Ala | Ala | Ala | Arg | Gly | Gly | Arg | Val | His | Arg | Ala | 325 | 330 | 335 |

| | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| Ala | Asp | Pro | Leu | Ala | Gly | Leu | Lys | Asp | Leu | Lys | Glu | Val | Arg | Gly | Leu | |
| | | | 340 | | | | | 345 | | | | | | 350 | | |
| Leu | Ala | Lys | Asp | Leu | Ala | Val | Leu | Ala | Ser | Arg | Glu | Gly | Leu | Asp | Leu | |
| | | 355 | | | | | 360 | | | | | 365 | | | | |
| Val | Pro | Gly | Asp | Asp | Pro | Met | Leu | Leu | Ala | Tyr | Leu | Leu | Asp | Pro | Ser | |
| | | 370 | | | | 375 | | | | | | 380 | | | | |
| Asn | Thr | Thr | Pro | Glu | Gly | Val | Ala | Arg | Arg | Tyr | Gly | Gly | Glu | Trp | Thr | |
| 385 | | | | | 390 | | | | | 395 | | | | | 400 | |
| Glu | Asp | Ala | Ala | His | Arg | Ala | Leu | Leu | Ser | Glu | Arg | Leu | His | Arg | Asn | |
| | | | | 405 | | | | | 410 | | | | | 415 | | |
| Leu | Leu | Lys | Arg | Leu | Glu | Gly | Glu | Glu | Lys | Leu | Leu | Trp | Leu | Tyr | His | |
| | | | 420 | | | | | 425 | | | | | 430 | | | |
| Glu | Val | Glu | Lys | Pro | Leu | Ser | Arg | Val | Leu | Ala | His | Met | Glu | Ala | Thr | |
| | | 435 | | | | | 440 | | | | | 445 | | | | |
| Gly | Val | Arg | Arg | Asp | Val | Ala | Tyr | Leu | Gln | Ala | Leu | Ser | Leu | Glu | Leu | |
| | 450 | | | | | 455 | | | | | 460 | | | | | |
| Ala | Glu | Glu | Ile | Arg | Arg | Leu | Glu | Glu | Glu | Val | Phe | Arg | Leu | Ala | Gly | |
| 465 | | | | | 470 | | | | | 475 | | | | | 480 | |
| His | Pro | Phe | Asn | Leu | Asn | Ser | Arg | Asp | Gln | Leu | Glu | Arg | Val | Leu | Phe | |
| | | | 485 | | | | | 490 | | | | | | 495 | | |
| Asp | Glu | Leu | Arg | Leu | Pro | Ala | Leu | Gly | Lys | Thr | Gln | Lys | Thr | Gly | Lys | |
| | | | 500 | | | | | 505 | | | | | | 510 | | |
| Arg | Ser | Thr | Ser | Ala | Ala | Val | Leu | Glu | Ala | Leu | Arg | Glu | Ala | His | Pro | |
| | | 515 | | | | | 520 | | | | | 525 | | | | |
| Ile | Val | Glu | Lys | Ile | Leu | Gln | His | Arg | Glu | Leu | Thr | Lys | Leu | Lys | Asn | |
| | 530 | | | | | 535 | | | | | 540 | | | | | |
| Thr | Tyr | Val | Asp | Pro | Leu | Pro | Ser | Leu | Val | His | Pro | Arg | Thr | Gly | Arg | |
| 545 | | | | | 550 | | | | | 555 | | | | | 560 | |
| Leu | His | Thr | Arg | Phe | Asn | Gln | Thr | Ala | Thr | Ala | Thr | Gly | Arg | Leu | Ser | |
| | | | | 565 | | | | 570 | | | | | | 575 | | |
| Ser | Ser | Asp | Pro | Asn | Leu | Gln | Asn | Ile | Pro | Val | Arg | Thr | Pro | Leu | Gly | |
| | | | 580 | | | | 585 | | | | | | 590 | | | |
| Gln | Arg | Ile | Arg | Arg | Ala | Phe | Ile | Ala | Glu | Glu | Gly | Trp | Leu | Leu | Val | |
| | | 595 | | | | | 600 | | | | | 605 | | | | |
| Ala | Leu | Asp | Tyr | Ser | Gln | Ile | Glu | Leu | Arg | Val | Leu | Ala | His | Leu | Ser | |
| | 610 | | | | | 615 | | | | | 620 | | | | | |
| Gly | Asp | Glu | Asn | Leu | Ile | Arg | Val | Phe | Gln | Glu | Gly | Arg | Asp | Ile | His | |
| 625 | | | | | 630 | | | | | 635 | | | | | 640 | |

Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
645 650 655

Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
660 665 670

Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
675 680 685

Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
690 695 700

Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
705 710 715 720

Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
725 730 735

Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
740 745 750

Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
755 760 765

Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
770 775 780

His Asp Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
785 790 795 800

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Glu

<210> 14
<211> 833
<212> PRT
<213> Artificial

<220>
<223> Synthetic

<400> 14

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1 5 10 15

Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
20 25 30

Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe

| | | |
|-----------------------------|-------------------------------------|-----|
| 340 | 345 | 350 |
| Leu Ala Lys Asp Leu Ser Val | Leu Ala Leu Arg Glu Gly Leu Gly Leu | |
| 355 | 360 | 365 |
| Pro Pro Gly Asp Asp Pro Met | Leu Leu Ala Tyr Leu Leu Asp Pro Ser | |
| 370 | 375 | 380 |
| Asn Thr Thr Pro Glu Gly Val | Ala Arg Arg Tyr Gly Gly Glu Trp Thr | |
| 385 | 390 | 395 |
| Glu Glu Ala Gly Glu Arg Ala | Ala Leu Ser Glu Arg Leu Phe Ala Asn | |
| 405 | 410 | 415 |
| Leu Trp Gly Arg Leu Glu Gly | Glu Glu Arg Leu Leu Trp Leu Tyr Arg | |
| 420 | 425 | 430 |
| Glu Val Glu Arg Pro Leu Ser | Ala Val Leu Ala His Met Glu Ala Thr | |
| 435 | 440 | 445 |
| Gly Val Arg Leu Asp Val Ala | Tyr Leu Arg Ala Leu Ser Leu Glu Val | |
| 450 | 455 | 460 |
| Ala Glu Glu Ile Ala Arg Leu | Glu Ala Glu Val Phe Arg Leu Ala Gly | |
| 465 | 470 | 475 |
| His Pro Phe Asn Leu Asn Ser | Arg Asp Gln Leu Glu Arg Val Leu Phe | |
| 485 | 490 | 495 |
| Asp Glu Leu Gly Leu Pro Ala | Ile Gly Lys Thr Glu Lys Thr Gly Lys | |
| 500 | 505 | 510 |
| Arg Ser Thr Ser Ala Ala Val | Leu Glu Ala Leu Arg Glu Ala His Pro | |
| 515 | 520 | 525 |
| Ile Val Glu Lys Ile Leu Gln | Tyr Arg Glu Leu Thr Lys Leu Lys Ser | |
| 530 | 535 | 540 |
| Thr Tyr Ile Asp Pro Leu Pro | Asp Leu Ile His Pro Arg Thr Gly Arg | |
| 545 | 550 | 555 |
| Leu His Thr Arg Phe Asn Gln | Thr Ala Thr Gly Arg Leu Ser | |
| 565 | 570 | 575 |
| Ser Ser Asp Pro Asn Leu Gln | Asn Ile Pro Val Arg Thr Pro Leu Gly | |
| 580 | 585 | 590 |
| Gln Arg Ile Arg Arg Ala Phe | Val Ala Glu Ala Gly Trp Ala Leu Val | |
| 595 | 600 | 605 |
| Ala Leu Asp Tyr Ser Gln Ile | Glu Leu Arg Val Leu Ala His Leu Ser | |
| 610 | 615 | 620 |
| Gly Asp Glu Asn Leu Ile Arg | Val Phe Gln Glu Gly Lys Asp Ile His | |
| 625 | 630 | 635 |
| Thr Gln Thr Ala Ser Trp Met | Phe Gly Val Pro Pro Glu Ala Val Asp | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| | 645 | | 650 | | 655 | | | | | | | | | | |
| Pro | Leu | Met | Arg | Arg | Ala | Ala | Lys | Thr | Val | Asn | Phe | Gly | Val | Leu | Tyr |
| | | | 660 | | | | | 665 | | | | | 670 | | |
| Gly | Met | Ser | Ala | His | Arg | Leu | Ser | Gln | Glu | Leu | Ala | Ile | Pro | Tyr | Glu |
| | | 675 | | | | | 680 | | | | | 685 | | | |
| Glu | Ala | Val | Ala | Phe | Ile | Glu | Arg | Tyr | Phe | Gln | Ser | Phe | Pro | Lys | Val |
| | 690 | | | | | 695 | | | | | 700 | | | | |
| Arg | Ala | Trp | Ile | Glu | Lys | Thr | Leu | Glu | Glu | Gly | Arg | Lys | Arg | Gly | Tyr |
| 705 | | | | | 710 | | | | | 715 | | | | | 720 |
| Val | Glu | Thr | Leu | Phe | Gly | Arg | Arg | Arg | Tyr | Val | Pro | Asp | Leu | Asn | Ala |
| | | | | 725 | | | | | 730 | | | | | 735 | |
| Arg | Val | Lys | Ser | Val | Arg | Glu | Ala | Ala | Glu | Arg | Met | Ala | Phe | Asn | Met |
| | | | 740 | | | | | 745 | | | | | 750 | | |
| Pro | Val | Gln | Gly | Thr | Ala | Ala | Asp | Leu | Met | Lys | Leu | Ala | Met | Val | Lys |
| | | 755 | | | | | 760 | | | | | 765 | | | |
| Leu | Phe | Pro | Arg | Leu | Arg | Glu | Met | Gly | Ala | Arg | Met | Leu | Leu | Gln | Val |
| | 770 | | | | | 775 | | | | | 780 | | | | |
| His | Asp | Glu | Leu | Leu | Leu | Glu | Ala | Pro | Gln | Ala | Arg | Ala | Glu | Glu | Val |
| 785 | | | | | 790 | | | | 795 | | | | | | 800 |
| Ala | Ala | Leu | Ala | Lys | Glu | Ala | Met | Glu | Lys | Ala | Tyr | Pro | Leu | Ala | Val |
| | | | | 805 | | | | | 810 | | | | | 815 | |
| Pro | Leu | Glu | Val | Glu | Val | Gly | Met | Gly | Glu | Asp | Trp | Leu | Ser | Ala | Lys |
| | | | 820 | | | | | 825 | | | | | 830 | | |

Gly

<210> 15
 <211> 833
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 15

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Asn | Ser | Gly | Met | Leu | Pro | Leu | Phe | Glu | Pro | Lys | Gly | Arg | Val | Leu |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Leu | Val | Asp | Gly | His | His | Leu | Ala | Tyr | Arg | Thr | Phe | His | Ala | Leu | Lys |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Gly | Leu | Thr | Thr | Ser | Arg | Gly | Glu | Pro | Val | Gln | Ala | Val | Tyr | Gly | Phe |
| | | 35 | | | | 40 | | | | | | 45 | | | |

Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
50 55 60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
65 70 75 80

Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
85 90 95

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
100 105 110

Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
115 120 125

Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
130 135 140

Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
145 150 155 160

Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
165 170 175

Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
180 185 190

Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
195 200 205

Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
210 215 220

Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
225 230 235 240

Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
245 250 255

Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
260 265 270

Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
275 280 285

Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu
290 295 300

Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala
305 310 315 320

Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala
325 330 335

Pro Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu
340 345 350

Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
660 665 670

Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
675 680 685

Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
690 695 700

Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
705 710 715 720

Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
725 730 735

Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
740 745 750

Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
755 760 765

Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
770 775 780

His Asp Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
785 790 795 800

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Glu

<210> 16
<211> 839
<212> PRT
<213> Artificial

<220>
<223> Synthetic

<400> 16

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1 5 10 15

Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
20 25 30

Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
35 40 45

Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
50 55 60

Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
675 680 685

Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
690 695 700

Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
705 710 715 720

Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
725 730 735

Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
740 745 750

Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
755 760 765

Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
770 775 780

His Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
785 790 795 800

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Glu His His His His His His
835

<210> 17
<211> 839
<212> PRT
<213> Artificial

<220>
<223> Synthetic

<400> 17

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1 5 10 15

Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
20 25 30

Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
35 40 45

Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
50 55 60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly

| | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Gly | Tyr | Lys | Ala | Gly | Arg | Ala | Pro | Thr | Pro | Glu | Asp | Phe | Pro | Arg | Gln | 85 | 90 | 95 |
| Leu | Ala | Leu | Ile | Lys | Glu | Leu | Val | Asp | Leu | Leu | Gly | Leu | Ala | Arg | Leu | 100 | 105 | 110 |
| Glu | Val | Pro | Gly | Tyr | Glu | Ala | Asp | Asp | Val | Leu | Ala | Ser | Leu | Ala | Lys | 115 | 120 | 125 |
| Lys | Ala | Glu | Lys | Glu | Gly | Tyr | Glu | Val | Arg | Ile | Leu | Thr | Ala | Asp | Lys | 130 | 135 | 140 |
| Asp | Leu | Tyr | Gln | Leu | Leu | Ser | Asp | Arg | Ile | His | Val | Leu | His | Pro | Glu | 145 | 150 | 155 |
| Gly | Tyr | Leu | Ile | Thr | Pro | Ala | Trp | Leu | Trp | Glu | Lys | Tyr | Gly | Leu | Arg | 165 | 170 | 175 |
| Pro | Asp | Gln | Trp | Ala | Asp | Tyr | Arg | Ala | Leu | Thr | Gly | Asp | Glu | Ser | Asp | 180 | 185 | 190 |
| Asn | Leu | Pro | Gly | Val | Lys | Gly | Ile | Gly | Glu | Lys | Thr | Ala | Arg | Lys | Leu | 195 | 200 | 205 |
| Leu | Glu | Glu | Trp | Gly | Ser | Leu | Glu | Ala | Leu | Leu | Lys | Asn | Leu | Asp | Arg | 210 | 215 | 220 |
| Leu | Lys | Pro | Ala | Ile | Arg | Glu | Lys | Ile | Leu | Ala | His | Met | Asp | Asp | Leu | 225 | 230 | 235 |
| Lys | Leu | Ser | Trp | Asp | Leu | Ala | Lys | Val | Arg | Thr | Asp | Leu | Pro | Leu | Glu | 245 | 250 | 255 |
| Val | Asp | Phe | Ala | Lys | Arg | Arg | Glu | Pro | Asp | Arg | Glu | Arg | Leu | Arg | Ala | 260 | 265 | 270 |
| Phe | Leu | Glu | Arg | Leu | Glu | Phe | Gly | Ser | Leu | Leu | His | Glu | Phe | Gly | Leu | 275 | 280 | 285 |
| Leu | Glu | Ser | Pro | Lys | Ala | Leu | Glu | Glu | Ala | Pro | Trp | Pro | Pro | Pro | Glu | 290 | 295 | 300 |
| Gly | Ala | Phe | Val | Gly | Phe | Val | Leu | Ser | Arg | Lys | Glu | Pro | Met | Trp | Ala | 305 | 310 | 315 |
| Asp | Leu | Leu | Ala | Leu | Ala | Ala | Ala | Arg | Gly | Gly | Arg | Val | His | Arg | Ala | 325 | 330 | 335 |
| Pro | Glu | Pro | Tyr | Lys | Ala | Leu | Arg | Asp | Leu | Lys | Glu | Ala | Arg | Gly | Leu | 340 | 345 | 350 |
| Leu | Ala | Lys | Asp | Leu | Ser | Val | Leu | Ala | Leu | Arg | Glu | Gly | Leu | Gly | Leu | 355 | 360 | 365 |
| Pro | Pro | Gly | Asp | Asp | Pro | Met | Leu | Leu | Ala | Tyr | Leu | Leu | Asp | Pro | Ser | 370 | 375 | 380 |

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
 100 105 110
 Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
 115 120 125
 Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
 130 135 140
 Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
 145 150 155 160
 Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
 165 170 175
 Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
 180 185 190
 Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
 195 200 205
 Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
 210 215 220
 Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
 225 230 235 240
 Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
 245 250 255
 Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
 260 265 270
 Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
 275 280 285
 Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu
 290 295 300
 Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala
 305 310 315 320
 Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala
 325 330 335
 Pro Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu
 340 345 350
 Leu Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu
 355 360 365
 Pro Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser
 370 375 380
 Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr
 385 390 395 400

| | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| Glu | Glu | Ala | Gly | Glu | Arg | Ala | Ala | Leu | Ser | Glu | Arg | Leu | Phe | Ala | Asn | |
| | | | | 405 | | | | | 410 | | | | | 415 | | |
| Leu | Trp | Gly | Arg | Leu | Glu | Gly | Glu | Glu | Arg | Leu | Leu | Trp | Leu | Tyr | Arg | |
| | | | 420 | | | | | 425 | | | | | 430 | | | |
| Glu | Val | Glu | Arg | Pro | Leu | Ser | Arg | Val | Leu | Ala | His | Met | Glu | Ala | Thr | |
| | | 435 | | | | | 440 | | | | | 445 | | | | |
| Gly | Val | Arg | Arg | Asp | Val | Ala | Tyr | Leu | Gln | Ala | Leu | Ser | Leu | Glu | Leu | |
| | 450 | | | | | 455 | | | | | 460 | | | | | |
| Ala | Glu | Glu | Ile | Arg | Arg | Leu | Glu | Glu | Glu | Val | Phe | Arg | Leu | Ala | Gly | |
| 465 | | | | | 470 | | | | | 475 | | | | | 480 | |
| His | Pro | Phe | Asn | Leu | Asn | Ser | Arg | Asp | Gln | Leu | Glu | Arg | Val | Leu | Phe | |
| | | | 485 | | | | | | 490 | | | | | 495 | | |
| Asp | Glu | Leu | Arg | Leu | Pro | Ala | Leu | Gly | Lys | Thr | Gln | Lys | Thr | Gly | Lys | |
| | | 500 | | | | | | 505 | | | | | 510 | | | |
| Arg | Ser | Thr | Ser | Ala | Ala | Val | Leu | Glu | Ala | Leu | Arg | Glu | Ala | His | Pro | |
| | | 515 | | | | | 520 | | | | | 525 | | | | |
| Ile | Val | Glu | Lys | Ile | Leu | Gln | His | Arg | Glu | Leu | Thr | Lys | Leu | Lys | Asn | |
| | 530 | | | | | 535 | | | | | 540 | | | | | |
| Thr | Tyr | Val | Asp | Pro | Leu | Pro | Ser | Leu | Val | His | Pro | Arg | Thr | Gly | Arg | |
| 545 | | | | | 550 | | | | | 555 | | | | | 560 | |
| Leu | His | Thr | Arg | Phe | Asn | Gln | Thr | Ala | Thr | Ala | Thr | Gly | Arg | Leu | Ser | |
| | | | | 565 | | | | | 570 | | | | | 575 | | |
| Ser | Ser | Asp | Pro | Asn | Leu | Gln | Asn | Ile | Pro | Val | Arg | Thr | Pro | Leu | Gly | |
| | | | 580 | | | | | 585 | | | | | 590 | | | |
| Gln | Arg | Ile | Arg | Arg | Ala | Phe | Ile | Ala | Glu | Glu | Gly | Trp | Leu | Leu | Val | |
| | | 595 | | | | | 600 | | | | | 605 | | | | |
| Ala | Leu | Asp | Tyr | Ser | Gln | Ile | Glu | Leu | Arg | Val | Leu | Ala | His | Leu | Ser | |
| | 610 | | | | | 615 | | | | | 620 | | | | | |
| Gly | Asp | Glu | Asn | Leu | Ile | Arg | Val | Phe | Gln | Glu | Gly | Arg | Asp | Ile | His | |
| 625 | | | | | 630 | | | | | 635 | | | | 640 | | |
| Thr | Glu | Thr | Ala | Ser | Trp | Met | Phe | Gly | Val | Pro | Arg | Glu | Ala | Val | Asp | |
| | | | | 645 | | | | | 650 | | | | | 655 | | |
| Pro | Leu | Met | Arg | Arg | Ala | Ala | Lys | Thr | Ile | Asn | Phe | Gly | Val | Leu | Tyr | |
| | | | 660 | | | | | 665 | | | | | 670 | | | |
| Gly | Met | Ser | Ala | His | Arg | Leu | Ser | Gln | Glu | Leu | Ala | Ile | Pro | Tyr | Glu | |
| | | 675 | | | | | 680 | | | | | 685 | | | | |
| Glu | Ala | Gln | Ala | Phe | Ile | Glu | Arg | Tyr | Phe | Gln | Ser | Phe | Pro | Lys | Val | |
| | 690 | | | | | 695 | | | | | 700 | | | | | |

Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
705 710 715 720

Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
725 730 735

Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
740 745 750

Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
755 760 765

Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
770 775 780

His Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
785 790 795 800

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Glu His His His His His His
835

<210> 20
<211> 839
<212> PRT
<213> Artificial

<220>
<223> Synthetic

<400> 20

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1 5 10 15

Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
20 25 30

Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
35 40 45

Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
50 55 60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
65 70 75 80

Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
85 90 95

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu

Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
 115 120 125
 Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
 130 135 140
 Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
 145 150 155 160
 Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
 165 170 175
 Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
 180 185 190
 Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
 195 200 205
 Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
 210 215 220
 Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
 225 230 235 240
 Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
 245 250 255
 Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
 260 265 270
 Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
 275 280 285
 Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu
 290 295 300
 Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala
 305 310 315 320
 Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala
 325 330 335
 Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val Arg Gly Leu
 340 345 350
 Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly Leu Asp Leu
 355 360 365
 Val Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser
 370 375 380
 Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr
 385 390 395 400
 Glu Asp Ala Ala His Arg Ala Leu Leu Ser Glu Arg Leu His Arg Asn

Leu Leu Lys Arg Leu Glu Gly Glu Glu Lys Leu Leu Trp Leu Tyr His
 420 425 430
 Glu Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met Glu Ala Thr
 435 440 445
 Gly Val Arg Leu Asp Val Ala Tyr Leu Gln Ala Leu Ser Leu Glu Val
 450 455 460
 Ala Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly
 465 470 475 480
 His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe
 485 490 495
 Asp Glu Leu Gly Leu Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys
 500 505 510
 Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro
 515 520 525
 Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser
 530 535 540
 Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg
 545 550 555 560
 Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser
 565 570 575
 Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly
 580 585 590
 Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val
 595 600 605
 Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser
 610 615 620
 Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His
 625 630 635 640
 Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
 645 650 655
 Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
 660 665 670
 Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
 675 680 685
 Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
 690 695 700
 Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
 705 710 715 720

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Glu | Val | Glu | Lys | Pro | Leu | Ser | Arg | Val | Leu | Ala | His | Met | Glu | Ala | Thr |
| | | 435 | | | | | 440 | | | | | 445 | | | |
| Gly | Val | Arg | Leu | Asp | Val | Ala | Tyr | Leu | Arg | Ala | Leu | Ser | Leu | Glu | Leu |
| | 450 | | | | | 455 | | | | | 460 | | | | |
| Ala | Glu | Glu | Ile | Ala | Arg | Leu | Glu | Ala | Glu | Val | Phe | Arg | Leu | Ala | Gly |
| 465 | | | | | 470 | | | | | 475 | | | | | 480 |
| His | Pro | Phe | Asn | Leu | Asn | Ser | Arg | Asp | Gln | Leu | Glu | Arg | Val | Leu | Phe |
| | | | 485 | | | | | | 490 | | | | | 495 | |
| Asp | Glu | Leu | Gly | Leu | Pro | Ala | Ile | Gly | Lys | Thr | Glu | Lys | Thr | Gly | Lys |
| | | | 500 | | | | | 505 | | | | | 510 | | |
| Arg | Ser | Thr | Ser | Ala | Ala | Val | Leu | Glu | Ala | Leu | Arg | Glu | Ala | His | Pro |
| | | 515 | | | | | 520 | | | | | 525 | | | |
| Ile | Val | Glu | Lys | Ile | Leu | Gln | Tyr | Arg | Glu | Leu | Thr | Lys | Leu | Lys | Ser |
| | 530 | | | | | 535 | | | | | 540 | | | | |
| Thr | Tyr | Ile | Asp | Pro | Leu | Pro | Asp | Leu | Ile | His | Pro | Arg | Thr | Gly | Arg |
| 545 | | | | | 550 | | | | | 555 | | | | | 560 |
| Leu | His | Thr | Arg | Phe | Asn | Gln | Thr | Ala | Thr | Ala | Thr | Gly | Arg | Leu | Ser |
| | | | | 565 | | | | | 570 | | | | | 575 | |
| Ser | Ser | Asp | Pro | Asn | Leu | Gln | Asn | Ile | Pro | Val | Arg | Thr | Pro | Leu | Gly |
| | | | 580 | | | | | 585 | | | | | 590 | | |
| Gln | Arg | Ile | Arg | Arg | Ala | Phe | Ile | Ala | Glu | Glu | Gly | Trp | Leu | Leu | Val |
| | | 595 | | | | | 600 | | | | | 605 | | | |
| Ala | Leu | Asp | Tyr | Ser | Gln | Ile | Glu | Leu | Arg | Val | Leu | Ala | His | Leu | Ser |
| | 610 | | | | | 615 | | | | | 620 | | | | |
| Gly | Asp | Glu | Asn | Leu | Ile | Arg | Val | Phe | Gln | Glu | Gly | Arg | Asp | Ile | His |
| 625 | | | | | 630 | | | | | 635 | | | | | 640 |
| Thr | Glu | Thr | Ala | Ser | Trp | Met | Phe | Gly | Val | Pro | Arg | Glu | Ala | Val | Asp |
| | | | | 645 | | | | | 650 | | | | | 655 | |
| Pro | Leu | Met | Arg | Arg | Ala | Ala | Lys | Thr | Ile | Asn | Phe | Gly | Val | Leu | Tyr |
| | | | 660 | | | | | 665 | | | | | 670 | | |
| Gly | Met | Ser | Ala | His | Arg | Leu | Ser | Gln | Glu | Leu | Ala | Ile | Pro | Tyr | Glu |
| | | 675 | | | | | 680 | | | | | 685 | | | |
| Glu | Ala | Gln | Ala | Phe | Ile | Glu | Arg | Tyr | Phe | Gln | Ser | Phe | Pro | Lys | Val |
| | 690 | | | | | 695 | | | | | 700 | | | | |
| Arg | Ala | Trp | Ile | Glu | Lys | Thr | Leu | Glu | Glu | Gly | Arg | Arg | Arg | Gly | Tyr |
| 705 | | | | | 710 | | | | | 715 | | | | | 720 |
| Val | Glu | Thr | Leu | Phe | Gly | Arg | Arg | Arg | Tyr | Val | Pro | Asp | Leu | Glu | Ala |
| | | | | 725 | | | | | 730 | | | | | 735 | |

Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
740 745 750

Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
755 760 765

Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
770 775 780

His Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
785 790 795 800

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Glu His His His His His His
835

<210> 23
<211> 839
<212> PRT
<213> Artificial

<220>
<223> Synthetic

<400> 23

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1 5 10 15

Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
20 25 30

Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
35 40 45

Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
50 55 60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
65 70 75 80

Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
85 90 95

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
100 105 110

Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
115 120 125

Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys

| | | | | | | | | | | | | | | | | |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|--|
| | 130 | | | | 135 | | | | 140 | | | | | | | |
| Asp 145 | Leu | Tyr | Gln | Leu | Leu | Ser | Asp | Arg | Ile | His | Val | Leu | His | Pro | Glu 160 | |
| Gly | Tyr | Leu | Ile | Thr 165 | Pro | Ala | Trp | Leu | Trp 170 | Glu | Lys | Tyr | Gly | Leu 175 | Arg | |
| Pro | Asp | Gln | Trp 180 | Ala | Asp | Tyr | Arg | Ala 185 | Leu | Thr | Gly | Asp | Glu 190 | Ser | Asp | |
| Asn | Leu | Pro 195 | Gly | Val | Lys | Gly | Ile 200 | Gly | Glu | Lys | Thr | Ala 205 | Arg | Lys | Leu | |
| Leu | Glu 210 | Glu | Trp | Gly | Ser | Leu | Glu 215 | Ala | Leu | Leu | Lys 220 | Asn | Leu | Asp | Arg | |
| Leu 225 | Lys | Pro | Ala | Ile | Arg 230 | Glu | Lys | Ile | Leu | Ala 235 | His | Met | Asp | Asp | Leu 240 | |
| Lys | Leu | Ser | Trp | Asp 245 | Leu | Ala | Lys | Val | Arg 250 | Thr | Asp | Leu | Pro | Leu 255 | Glu | |
| Val | Asp | Phe | Ala 260 | Lys | Arg | Arg | Glu | Pro 265 | Asp | Arg | Glu | Arg | Leu 270 | Arg | Ala | |
| Phe | Leu | Glu 275 | Arg | Leu | Glu | Phe | Gly 280 | Ser | Leu | Leu | His | Glu 285 | Phe | Gly | Leu | |
| Leu 290 | Glu | Ser | Pro | Lys | Ala | Leu | Glu 295 | Glu | Ala | Pro | Trp 300 | Pro | Pro | Pro | Glu | |
| Gly 305 | Ala | Phe | Val | Gly | Phe 310 | Val | Leu | Ser | Arg | Lys 315 | Glu | Pro | Met | Trp | Ala 320 | |
| Asp | Leu | Leu | Ala | Leu 325 | Ala | Ala | Ala | Arg | Gly 330 | Gly | Arg | Val | His | Arg 335 | Ala | |
| Ala | Asp | Pro | Leu 340 | Ala | Gly | Leu | Lys | Asp 345 | Leu | Lys | Glu | Val | Arg 350 | Gly | Leu | |
| Leu | Ala | Lys 355 | Asp | Leu | Ala | Val | Leu 360 | Ala | Ser | Arg | Glu | Gly 365 | Leu | Asp | Leu | |
| Val 370 | Pro | Gly | Asp | Asp | Pro | Met 375 | Leu | Leu | Ala | Tyr | Leu 380 | Leu | Asp | Pro | Ser | |
| Asn 385 | Thr | Thr | Pro | Glu | Gly 390 | Val | Ala | Arg | Arg | Tyr 395 | Gly | Gly | Glu | Trp | Thr 400 | |
| Glu | Asp | Ala | Ala | His 405 | Arg | Ala | Leu | Leu | Ser 410 | Glu | Arg | Leu | His | Arg 415 | Asn | |
| Leu | Leu | Lys 420 | Arg | Leu | Glu | Gly | Glu 425 | Glu | Lys | Leu | Leu | Trp | Leu 430 | Tyr | His | |
| Glu | Val | Glu | Lys | Pro | Leu | Ser | Arg | Val | Leu | Ala | His | Met | Glu | Ala | Thr | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| | 435 | | | | | 440 | | | | | 445 | | | | |
| Gly | Val | Arg | Leu | Asp | Val | Ala | Tyr | Leu | Arg | Ala | Leu | Ser | Leu | Glu | Val |
| | 450 | | | | | 455 | | | | | 460 | | | | |
| Ala | Glu | Glu | Ile | Arg | Arg | Leu | Glu | Ala | Glu | Val | Phe | Arg | Leu | Ala | Gly |
| 465 | | | | | 470 | | | | | 475 | | | | | 480 |
| His | Pro | Phe | Asn | Leu | Asn | Ser | Arg | Asp | Gln | Leu | Glu | Arg | Val | Leu | Phe |
| | | | | 485 | | | | | 490 | | | | | 495 | |
| Asp | Glu | Leu | Gly | Leu | Pro | Ala | Ile | Gly | Lys | Thr | Glu | Lys | Thr | Gly | Lys |
| | | | 500 | | | | | 505 | | | | | 510 | | |
| Arg | Ser | Thr | Ser | Ala | Ala | Val | Leu | Glu | Ala | Leu | Arg | Glu | Ala | His | Pro |
| | | 515 | | | | | 520 | | | | | 525 | | | |
| Ile | Val | Glu | Lys | Ile | Leu | Gln | Tyr | Arg | Glu | Leu | Thr | Lys | Leu | Lys | Ser |
| | 530 | | | | | 535 | | | | | 540 | | | | |
| Thr | Tyr | Ile | Asp | Pro | Leu | Pro | Asp | Leu | Ile | His | Pro | Arg | Thr | Gly | Arg |
| 545 | | | | | 550 | | | | | 555 | | | | | 560 |
| Leu | His | Thr | Arg | Phe | Asn | Gln | Thr | Ala | Thr | Ala | Thr | Gly | Arg | Leu | Ser |
| | | | | 565 | | | | | 570 | | | | | 575 | |
| Ser | Ser | Asp | Pro | Asn | Leu | Gln | Asn | Ile | Pro | Val | Arg | Thr | Pro | Leu | Gly |
| | | | 580 | | | | | 585 | | | | | 590 | | |
| Gln | Arg | Ile | Arg | Arg | Ala | Phe | Ile | Ala | Glu | Glu | Gly | Trp | Leu | Leu | Val |
| | | 595 | | | | | 600 | | | | | 605 | | | |
| Ala | Leu | Asp | Tyr | Ser | Gln | Ile | Glu | Leu | Arg | Val | Leu | Ala | His | Leu | Ser |
| | 610 | | | | | 615 | | | | | 620 | | | | |
| Gly | Asp | Glu | Asn | Leu | Ile | Arg | Val | Phe | Gln | Glu | Gly | Arg | Asp | Ile | His |
| 625 | | | | | 630 | | | | | 635 | | | | | 640 |
| Thr | Glu | Thr | Ala | Ser | Trp | Met | Phe | Gly | Val | Pro | Arg | Glu | Ala | Val | Asp |
| | | | | 645 | | | | | 650 | | | | | 655 | |
| Pro | Leu | Met | Arg | Arg | Ala | Ala | Lys | Thr | Ile | Asn | Phe | Gly | Val | Leu | Tyr |
| | | | 660 | | | | | 665 | | | | | 670 | | |
| Gly | Met | Ser | Ala | His | Arg | Leu | Ser | Gln | Glu | Leu | Ala | Ile | Pro | Tyr | Glu |
| | | 675 | | | | | 680 | | | | | 685 | | | |
| Glu | Ala | Gln | Ala | Phe | Ile | Glu | Arg | Tyr | Phe | Gln | Ser | Phe | Pro | Lys | Val |
| | 690 | | | | | 695 | | | | | 700 | | | | |
| Arg | Ala | Trp | Ile | Glu | Lys | Thr | Leu | Glu | Glu | Gly | Arg | Arg | Arg | Gly | Tyr |
| 705 | | | | | 710 | | | | | 715 | | | | | 720 |
| Val | Glu | Thr | Leu | Phe | Gly | Arg | Arg | Arg | Tyr | Val | Pro | Asp | Leu | Glu | Ala |
| | | | | 725 | | | | | 730 | | | | | 735 | |
| Arg | Val | Lys | Ser | Val | Arg | Glu | Ala | Ala | Glu | Arg | Met | Ala | Phe | Asn | Met |

```

740              745              750
Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
    755                    760                765

Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
    770                    775                780

His Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
    785                    790                795                800

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
            805                    810                815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
        820                    825                830

Glu His His His His His His
    835

<210>   24
<211>   839
<212>   PRT
<213> Artificial

<220>
<223> Synthetic

<400>   24


Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1          5          10           15

Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
      20                      25                 30

Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
     35                     40                   45

Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
     50                       55               60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
65         70           75             80

Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
       85                  90

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
      100                    105                110

Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
      115                    120                125

Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
      130                    135                140
```


| | | | |
|---|-----|-----|-----|
| Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg | 165 | 170 | 175 |
| Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp | 180 | 185 | 190 |
| Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu | 195 | 200 | 205 |
| Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg | 210 | 215 | 220 |
| Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu | 225 | 230 | 235 |
| Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu | 245 | 250 | 255 |
| Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala | 260 | 265 | 270 |
| Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu | 275 | 280 | 285 |
| Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu | 290 | 295 | 300 |
| Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala | 305 | 310 | 315 |
| Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala | 325 | 330 | 335 |
| Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val Arg Gly Leu | 340 | 345 | 350 |
| Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly Leu Asp Leu | 355 | 360 | 365 |
| Val Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser | 370 | 375 | 380 |
| Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr | 385 | 390 | 395 |
| Glu Asp Ala Ala His Arg Ala Leu Leu Ser Glu Arg Leu His Arg Asn | 405 | 410 | 415 |
| Leu Leu Lys Arg Leu Glu Gly Glu Glu Lys Leu Leu Trp Leu Tyr His | 420 | 425 | 430 |
| Glu Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met Glu Ala Thr | 435 | 440 | 445 |
| Gly Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val | 450 | 455 | 460 |

| | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| Ala | Glu | Glu | Ile | Ala | Arg | Leu | Glu | Ala | Glu | Val | Phe | Arg | Leu | Ala | Gly | |
| 465 | | | | | 470 | | | | | 475 | | | | | 480 | |
| His | Pro | Phe | Asn | Leu | Asn | Ser | Arg | Asp | Gln | Leu | Glu | Arg | Val | Leu | Phe | |
| | | | 485 | | | | | | 490 | | | | | 495 | | |
| Asp | Glu | Leu | Arg | Leu | Pro | Ala | Ile | Gly | Lys | Thr | Glu | Lys | Thr | Gly | Lys | |
| | | 500 | | | | | | 505 | | | | | | 510 | | |
| Arg | Ser | Thr | Ser | Ala | Ala | Val | Leu | Glu | Ala | Leu | Arg | Glu | Ala | His | Pro | |
| | | 515 | | | | | 520 | | | | | 525 | | | | |
| Ile | Val | Glu | Lys | Ile | Leu | Gln | Tyr | Arg | Glu | Leu | Thr | Lys | Leu | Lys | Ser | |
| | 530 | | | | | 535 | | | | | 540 | | | | | |
| Thr | Tyr | Ile | Asp | Pro | Leu | Pro | Asp | Leu | Ile | His | Pro | Arg | Thr | Gly | Arg | |
| 545 | | | | | 550 | | | | | 555 | | | | | 560 | |
| Leu | His | Thr | Arg | Phe | Asn | Gln | Thr | Ala | Thr | Ala | Thr | Gly | Arg | Leu | Ser | |
| | | | | 565 | | | | 570 | | | | | | 575 | | |
| Ser | Ser | Asp | Pro | Asn | Leu | Gln | Asn | Ile | Pro | Val | Arg | Thr | Pro | Leu | Gly | |
| | | 580 | | | | | | 585 | | | | | 590 | | | |
| Gln | Arg | Ile | Arg | Arg | Ala | Phe | Ile | Ala | Glu | Glu | Gly | Trp | Leu | Leu | Val | |
| | | 595 | | | | | 600 | | | | | 605 | | | | |
| Ala | Leu | Asp | Tyr | Ser | Gln | Ile | Glu | Leu | Arg | Val | Leu | Ala | His | Leu | Ser | |
| | 610 | | | | | 615 | | | | | 620 | | | | | |
| Gly | Asp | Glu | Asn | Leu | Ile | Arg | Val | Phe | Gln | Glu | Gly | Arg | Asp | Ile | His | |
| 625 | | | | 630 | | | | | | 635 | | | | | 640 | |
| Thr | Glu | Thr | Ala | Ser | Trp | Met | Phe | Gly | Val | Pro | Arg | Glu | Ala | Val | Asp | |
| | | | 645 | | | | | | 650 | | | | | 655 | | |
| Pro | Leu | Met | Arg | Arg | Ala | Ala | Lys | Thr | Ile | Asn | Phe | Gly | Val | Leu | Tyr | |
| | | 660 | | | | | | 665 | | | | | 670 | | | |
| Gly | Met | Ser | Ala | His | Arg | Leu | Ser | Gln | Glu | Leu | Ala | Ile | Pro | Tyr | Glu | |
| | 675 | | | | | 680 | | | | | | 685 | | | | |
| Glu | Ala | Gln | Ala | Phe | Ile | Glu | Arg | Tyr | Phe | Gln | Ser | Phe | Pro | Lys | Val | |
| | 690 | | | | | 695 | | | | | 700 | | | | | |
| Arg | Ala | Trp | Ile | Glu | Lys | Thr | Leu | Glu | Glu | Gly | Arg | Arg | Arg | Gly | Tyr | |
| 705 | | | | | 710 | | | | | 715 | | | | | 720 | |
| Val | Glu | Thr | Leu | Phe | Gly | Arg | Arg | Arg | Tyr | Val | Pro | Asp | Leu | Glu | Ala | |
| | | | 725 | | | | | | 730 | | | | | 735 | | |
| Arg | Val | Lys | Ser | Val | Arg | Glu | Ala | Ala | Glu | Arg | Met | Ala | Phe | Asn | Met | |
| | | 740 | | | | | | 745 | | | | | 750 | | | |
| Pro | Val | Gln | Gly | Thr | Ala | Ala | Asp | Leu | Met | Lys | Leu | Ala | Met | Val | Lys | |
| | | 755 | | | | | 760 | | | | | 765 | | | | |

| | | |
|---|-----|-----|
| 165 | 170 | 175 |
| Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp | | |
| 180 | 185 | 190 |
| Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu | | |
| 195 | 200 | 205 |
| Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg | | |
| 210 | 215 | 220 |
| Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu | | |
| 225 | 230 | 235 |
| Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu | | |
| 245 | 250 | 255 |
| Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala | | |
| 260 | 265 | 270 |
| Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu | | |
| 275 | 280 | 285 |
| Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu | | |
| 290 | 295 | 300 |
| Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala | | |
| 305 | 310 | 315 |
| Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala | | |
| 325 | 330 | 335 |
| Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val Arg Gly Leu | | |
| 340 | 345 | 350 |
| Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly Leu Asp Leu | | |
| 355 | 360 | 365 |
| Val Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser | | |
| 370 | 375 | 380 |
| Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr | | |
| 385 | 390 | 395 |
| Glu Asp Ala Ala His Arg Ala Leu Leu Ser Glu Arg Leu His Arg Asn | | |
| 405 | 410 | 415 |
| Leu Leu Lys Arg Leu Glu Gly Glu Glu Lys Leu Leu Trp Leu Tyr His | | |
| 420 | 425 | 430 |
| Glu Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met Glu Ala Thr | | |
| 435 | 440 | 445 |
| Gly Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val | | |
| 450 | 455 | 460 |
| Ala Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly | | |

| | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| His | Pro | Phe | Asn | Leu | Asn | Ser | Arg | Asp | Gln | Leu | Glu | Arg | Val | Leu | Phe | 485 | 490 | 495 |
| Asp | Glu | Leu | Gly | Leu | Pro | Ala | Ile | Gly | Lys | Thr | Glu | Lys | Thr | Gly | Lys | 500 | 505 | 510 |
| Arg | Ser | Thr | Ser | Ala | Ala | Val | Leu | Glu | Ala | Leu | Arg | Glu | Ala | His | Pro | 515 | 520 | 525 |
| Ile | Val | Glu | Lys | Ile | Leu | Gln | His | Arg | Glu | Leu | Thr | Lys | Leu | Lys | Ser | 530 | 535 | 540 |
| Thr | Tyr | Ile | Asp | Pro | Leu | Pro | Asp | Leu | Ile | His | Pro | Arg | Thr | Gly | Arg | 545 | 550 | 555 |
| Leu | His | Thr | Arg | Phe | Asn | Gln | Thr | Ala | Thr | Ala | Thr | Gly | Arg | Leu | Ser | 565 | 570 | 575 |
| Ser | Ser | Asp | Pro | Asn | Leu | Gln | Asn | Ile | Pro | Val | Arg | Thr | Pro | Leu | Gly | 580 | 585 | 590 |
| Gln | Arg | Ile | Arg | Arg | Ala | Phe | Ile | Ala | Glu | Glu | Gly | Trp | Leu | Leu | Val | 595 | 600 | 605 |
| Ala | Leu | Asp | Tyr | Ser | Gln | Ile | Glu | Leu | Arg | Val | Leu | Ala | His | Leu | Ser | 610 | 615 | 620 |
| Gly | Asp | Glu | Asn | Leu | Ile | Arg | Val | Phe | Gln | Glu | Gly | Arg | Asp | Ile | His | 625 | 630 | 635 |
| Thr | Glu | Thr | Ala | Ser | Trp | Met | Phe | Gly | Val | Pro | Arg | Glu | Ala | Val | Asp | 645 | 650 | 655 |
| Pro | Leu | Met | Arg | Arg | Ala | Ala | Lys | Thr | Ile | Asn | Phe | Gly | Val | Leu | Tyr | 660 | 665 | 670 |
| Gly | Met | Ser | Ala | His | Arg | Leu | Ser | Gln | Glu | Leu | Ala | Ile | Pro | Tyr | Glu | 675 | 680 | 685 |
| Glu | Ala | Gln | Ala | Phe | Ile | Glu | Arg | Tyr | Phe | Gln | Ser | Phe | Pro | Lys | Val | 690 | 695 | 700 |
| Arg | Ala | Trp | Ile | Glu | Lys | Thr | Leu | Glu | Glu | Gly | Arg | Arg | Arg | Gly | Tyr | 705 | 710 | 715 |
| Val | Glu | Thr | Leu | Phe | Gly | Arg | Arg | Arg | Tyr | Val | Pro | Asp | Leu | Glu | Ala | 725 | 730 | 735 |
| Arg | Val | Lys | Ser | Val | Arg | Glu | Ala | Ala | Glu | Arg | Met | Ala | Phe | Asn | Met | 740 | 745 | 750 |
| Pro | Val | Gln | Gly | Thr | Ala | Ala | Asp | Leu | Met | Lys | Leu | Ala | Met | Val | Lys | 755 | 760 | 765 |
| Leu | Phe | Pro | Arg | Leu | Glu | Glu | Met | Gly | Ala | Arg | Met | Leu | Leu | Gln | Val | 770 | 775 | 780 |

His Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
 785 790 795 800

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
 805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
 820 825 830

Glu His His His His His His 835

<210> 28

<211> 839

<212> PRT

<213> Artificial

<220>

<223> Synthetic

<400> 28

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
 1 5 10 15

Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
 20 25 30

Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
 35 40 45

Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
 50 55 60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
 65 70 75 80

Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
 85 90 95

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
 100 105 110

Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
 115 120 125

Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
 130 135 140

Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
 145 150 155 160

Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
 165 170 175

Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
 180 185 190

104230 924350

| | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|--|
| Asp | Glu | Leu | Gly | Leu | Pro | Ala | Ile | Gly | Lys | Thr | Glu | Lys | Thr | Gly | Lys | | |
| | | | 500 | | | | | | 505 | | | | | 510 | | | |
| Arg | Ser | Thr | Ser | Ala | Ala | Val | Leu | Glu | Ala | Leu | Arg | Glu | Ala | His | Pro | | |
| | | 515 | | | | | 520 | | | | | 525 | | | | | |
| Ile | Val | Glu | Lys | Ile | Leu | Gln | Tyr | Arg | Glu | Leu | Thr | Lys | Leu | Lys | Asn | | |
| | 530 | | | | | 535 | | | | | 540 | | | | | | |
| Thr | Tyr | Ile | Asp | Pro | Leu | Pro | Asp | Leu | Ile | His | Pro | Arg | Thr | Gly | Arg | | |
| 545 | | | | | 550 | | | | | 555 | | | | | 560 | | |
| Leu | His | Thr | Arg | Phe | Asn | Gln | Thr | Ala | Thr | Ala | Thr | Gly | Arg | Leu | Ser | | |
| | | | | 565 | | | | | 570 | | | | | 575 | | | |
| Ser | Ser | Asp | Pro | Asn | Leu | Gln | Asn | Ile | Pro | Val | Arg | Thr | Pro | Leu | Gly | | |
| | | | 580 | | | | | 585 | | | | | 590 | | | | |
| Gln | Arg | Ile | Arg | Arg | Ala | Phe | Ile | Ala | Glu | Glu | Gly | Trp | Leu | Leu | Val | | |
| | | 595 | | | | | 600 | | | | | 605 | | | | | |
| Ala | Leu | Asp | Tyr | Ser | Gln | Ile | Glu | Leu | Arg | Val | Leu | Ala | His | Leu | Ser | | |
| | 610 | | | | | 615 | | | | | 620 | | | | | | |
| Gly | Asp | Glu | Asn | Leu | Ile | Arg | Val | Phe | Gln | Glu | Gly | Arg | Asp | Ile | His | | |
| 625 | | | | 630 | | | | | | 635 | | | | 640 | | | |
| Thr | Glu | Thr | Ala | Ser | Trp | Met | Phe | Gly | Val | Pro | Arg | Glu | Ala | Val | Asp | | |
| | | | | 645 | | | | | 650 | | | | | 655 | | | |
| Pro | Leu | Met | Arg | Arg | Ala | Ala | Lys | Thr | Ile | Asn | Phe | Gly | Val | Leu | Tyr | | |
| | | | 660 | | | | | 665 | | | | | 670 | | | | |
| Gly | Met | Ser | Ala | His | Arg | Leu | Ser | Gln | Glu | Leu | Ala | Ile | Pro | Tyr | Glu | | |
| | 675 | | | | | 680 | | | | | | 685 | | | | | |
| Glu | Ala | Gln | Ala | Phe | Ile | Glu | Arg | Tyr | Phe | Gln | Ser | Phe | Pro | Lys | Val | | |
| | 690 | | | | | 695 | | | | | 700 | | | | | | |
| Arg | Ala | Trp | Ile | Glu | Lys | Thr | Leu | Glu | Glu | Gly | Arg | Arg | Arg | Gly | Tyr | | |
| 705 | | | | | 710 | | | | | 715 | | | | | 720 | | |
| Val | Glu | Thr | Leu | Phe | Gly | Arg | Arg | Arg | Tyr | Val | Pro | Asp | Leu | Glu | Ala | | |
| | | | | 725 | | | | | 730 | | | | | 735 | | | |
| Arg | Val | Lys | Ser | Val | Arg | Glu | Ala | Ala | Glu | Arg | Met | Ala | Phe | Asn | Met | | |
| | | | 740 | | | | | 745 | | | | | 750 | | | | |
| Pro | Val | Gln | Gly | Thr | Ala | Ala | Asp | Leu | Met | Lys | Leu | Ala | Met | Val | Lys | | |
| | | 755 | | | | | 760 | | | | | 765 | | | | | |
| Leu | Phe | Pro | Arg | Leu | Glu | Glu | Met | Gly | Ala | Arg | Met | Leu | Leu | Gln | Val | | |
| | 770 | | | | | 775 | | | | | 780 | | | | | | |
| His | Asn | Glu | Leu | Val | Leu | Glu | Ala | Pro | Lys | Glu | Arg | Ala | Glu | Ala | Val | | |
| 785 | | | | | 790 | | | | | 795 | | | | | 800 | | |

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Glu His His His His His His
835

<210> 29

<211> 839

<212> PRT

<213> Artificial

<220>

<223> Synthetic

<400> 29

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1 5 10 15

Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
20 25 30

Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
35 40 45

Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
50 55 60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
65 70 75 80

Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
85 90 95

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
100 105 110

Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
115 120 125

Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
130 135 140

Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
145 150 155 160

Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
165 170 175

Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
180 185 190

Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
195 200 205

09364636 052404

[illegible]

830

| | |
|-------|------------|
| <210> | 31 |
| <211> | 839 |
| <212> | PRT |
| <213> | Artificial |
| <220> | |
| <223> | Synthetic |
| <400> | 31 |

| | | | | | | | | | | | | | | | |
|------------|-----------|------------|------------|-----------|------------|------------|------------|------------|------------|------------|------------|------------|------------|-----------|------------|
| Met 1 | Asn | Ser | Gly | Met 5 | Leu | Pro | Leu | Phe | Glu 10 | Pro | Lys | Gly | Arg | Val 15 | Leu |
| Leu | Val | Asp | Gly 20 | His | His | Leu | Ala | Tyr 25 | Arg | Thr | Phe | His | Ala 30 | Leu | Lys |
| Gly | Leu | Thr 35 | Thr | Ser | Arg | Gly | Glu 40 | Pro | Val | Gln | Ala | Val 45 | Tyr | Gly | Phe |
| Ala | Lys 50 | Ser | Leu | Leu | Lys | Ala 55 | Leu | Lys | Glu | Asp | Gly 60 | Asp | Ala | Val | Ile |
| Val 65 | Val | Phe | Asp | Ala | Lys 70 | Ala | Pro | Ser | Phe | Arg 75 | His | Glu | Ala | Tyr | Gly 80 |
| Gly | Tyr | Lys | Ala | Gly 85 | Arg | Ala | Pro | Thr | Pro 90 | Glu | Asp | Phe | Pro | Arg 95 | Gln |
| Leu | Ala | Leu | Ile 100 | Lys | Glu | Leu | Val | Asp 105 | Leu | Leu | Gly | Leu | Ala 110 | Arg | Leu |
| Glu | Val | Pro 115 | Gly | Tyr | Glu | Ala | Asp 120 | Asp | Val | Leu | Ala | Ser 125 | Leu | Ala | Lys |
| Lys 130 | Ala | Glu | Lys | Glu | Gly | Tyr 135 | Glu | Val | Arg | Ile | Leu 140 | Thr | Ala | Asp | Lys |
| Asp 145 | Leu | Tyr | Gln | Leu | Leu 150 | Ser | Asp | Arg | Ile | His 155 | Val | Leu | His | Pro | Glu 160 |
| Gly | Tyr | Leu | Ile 165 | Thr | Pro | Ala | Trp | Leu | Trp 170 | Glu | Lys | Tyr | Gly | Leu | Arg |
| Pro | Asp | Gln 180 | Trp | Ala | Asp | Tyr | Arg | Ala 185 | Leu | Thr | Gly | Asp | Glu 190 | Ser | Asp |
| Asn | Leu | Pro 195 | Gly | Val | Lys | Gly | Ile 200 | Gly | Glu | Lys | Thr | Ala 205 | Arg | Lys | Leu |
| Leu 210 | Glu | Glu | Trp | Gly | Ser | Leu 215 | Glu | Ala | Leu | Leu | Lys 220 | Asn | Leu | Asp | Arg |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ile | Val | Glu | Lys | Ile | Leu | Gln | Tyr | Arg | Glu | Leu | Thr | Lys | Leu | Lys | Ser |
| 530 | | | | | | 535 | | | | | | 540 | | | |
| Thr | Tyr | Ile | Asp | Pro | Leu | Pro | Asp | Leu | Ile | His | Pro | Arg | Thr | Gly | Arg |
| 545 | | | | | 550 | | | | | 555 | | | | | 560 |
| Leu | His | Thr | Arg | Phe | Asn | Gln | Thr | Ala | Thr | Ala | Thr | Gly | Arg | Leu | Ser |
| | | | | 565 | | | | | 570 | | | | | 575 | |
| Ser | Ser | Asp | Pro | Asn | Leu | Gln | Asn | Ile | Pro | Val | Arg | Thr | Pro | Leu | Gly |
| | | | 580 | | | | | 585 | | | | | 590 | | |
| Gln | Arg | Ile | Arg | Arg | Ala | Phe | Ile | Ala | Glu | Glu | Gly | Trp | Leu | Leu | Val |
| | | 595 | | | | | 600 | | | | | 605 | | | |
| Ala | Leu | Asp | Tyr | Ser | Gln | Ile | Glu | Leu | Arg | Val | Leu | Ala | His | Leu | Ser |
| | 610 | | | | | 615 | | | | | 620 | | | | |
| Gly | Asp | Glu | Asn | Leu | Ile | Arg | Val | Phe | Gln | Glu | Gly | Arg | Asp | Ile | His |
| 625 | | | | | 630 | | | | | 635 | | | | | 640 |
| Thr | Glu | Thr | Ala | Ser | Trp | Met | Phe | Gly | Val | Pro | Arg | Glu | Ala | Val | Asp |
| | | | | 645 | | | | | 650 | | | | | 655 | |
| Pro | Leu | Met | Arg | Arg | Ala | Ala | Lys | Thr | Ile | Asn | Phe | Gly | Val | Leu | Tyr |
| | | | 660 | | | | | 665 | | | | | 670 | | |
| Gly | Met | Ser | Ala | His | Arg | Leu | Ser | Gln | Glu | Leu | Ala | Ile | Pro | Tyr | Glu |
| | | 675 | | | | | 680 | | | | | 685 | | | |
| Glu | Ala | Gln | Ala | Phe | Ile | Glu | Arg | Tyr | Phe | Gln | Ser | Phe | Pro | Lys | Val |
| | 690 | | | | | 695 | | | | | 700 | | | | |
| Arg | Ala | Trp | Ile | Glu | Lys | Thr | Leu | Glu | Glu | Gly | Arg | Arg | Arg | Gly | Tyr |
| 705 | | | | | 710 | | | | | 715 | | | | | 720 |
| Val | Glu | Thr | Leu | Phe | Gly | Arg | Arg | Arg | Tyr | Val | Pro | Asp | Leu | Glu | Ala |
| | | | | 725 | | | | | 730 | | | | | 735 | |
| Arg | Val | Lys | Ser | Val | Arg | Glu | Ala | Ala | Glu | Arg | Met | Ala | Phe | Asn | Met |
| | | | 740 | | | | | 745 | | | | | 750 | | |
| Pro | Val | Gln | Gly | Thr | Ala | Ala | Asp | Leu | Met | Lys | Leu | Ala | Met | Val | Lys |
| | | 755 | | | | | 760 | | | | | 765 | | | |
| Leu | Phe | Pro | Arg | Leu | Glu | Glu | Met | Gly | Ala | Arg | Met | Leu | Leu | Gln | Val |
| | 770 | | | | | 775 | | | | | 780 | | | | |
| His | Asn | Glu | Leu | Val | Leu | Glu | Ala | Pro | Lys | Glu | Arg | Ala | Glu | Ala | Val |
| 785 | | | | | 790 | | | | | 795 | | | | | 800 |
| Ala | Arg | Leu | Ala | Lys | Glu | Val | Met | Glu | Gly | Val | Tyr | Pro | Leu | Ala | Val |
| | | | | 805 | | | | | 810 | | | | | 815 | |
| Pro | Leu | Glu | Val | Glu | Val | Gly | Ile | Gly | Glu | Asp | Trp | Leu | Ser | Ala | Lys |
| | | | 820 | | | | | 825 | | | | | 830 | | |

| | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Lys | Leu | Ser | Trp | Asp | Leu | Ala | Lys | Val | Arg | Thr | Asp | Leu | Pro | Leu | Glu | 245 | 250 | 255 |
| Val | Asp | Phe | Ala | Lys | Arg | Arg | Glu | Pro | Asp | Arg | Glu | Arg | Leu | Arg | Ala | 260 | 265 | 270 |
| Phe | Leu | Glu | Arg | Leu | Glu | Phe | Gly | Ser | Leu | Leu | His | Glu | Phe | Gly | Leu | 275 | 280 | 285 |
| Leu | Glu | Ser | Pro | Lys | Ala | Leu | Glu | Glu | Ala | Pro | Trp | Pro | Pro | Pro | Glu | 290 | 295 | 300 |
| Gly | Ala | Phe | Val | Gly | Phe | Val | Leu | Ser | Arg | Lys | Glu | Pro | Met | Trp | Ala | 305 | 310 | 315 |
| Asp | Leu | Leu | Ala | Leu | Ala | Ala | Ala | Arg | Gly | Gly | Arg | Val | His | Arg | Ala | 325 | 330 | 335 |
| Pro | Glu | Pro | Tyr | Lys | Ala | Leu | Arg | Asp | Leu | Lys | Glu | Ala | Arg | Gly | Leu | 340 | 345 | 350 |
| Leu | Ala | Lys | Asp | Leu | Ser | Val | Leu | Ala | Leu | Arg | Glu | Gly | Leu | Gly | Leu | 355 | 360 | 365 |
| Pro | Pro | Gly | Asp | Asp | Pro | Met | Leu | Leu | Ala | Tyr | Leu | Leu | Asp | Pro | Ser | 370 | 375 | 380 |
| Asn | Thr | Thr | Pro | Glu | Gly | Val | Ala | Arg | Arg | Tyr | Gly | Gly | Glu | Trp | Thr | 385 | 390 | 395 |
| Glu | Glu | Ala | Gly | Glu | Arg | Ala | Ala | Leu | Ser | Glu | Arg | Leu | Phe | Ala | Asn | 405 | 410 | 415 |
| Leu | Leu | Gly | Arg | Leu | Glu | Gly | Glu | Glu | Arg | Leu | Leu | Trp | Leu | Tyr | Arg | 420 | 425 | 430 |
| Glu | Val | Glu | Arg | Pro | Leu | Ser | Ala | Val | Leu | Ala | His | Met | Glu | Ala | Thr | 435 | 440 | 445 |
| Gly | Val | Arg | Leu | Asp | Val | Ala | Tyr | Leu | Arg | Ala | Leu | Ser | Leu | Glu | Val | 450 | 455 | 460 |
| Ala | Glu | Glu | Ile | Ala | Arg | Leu | Glu | Ala | Glu | Val | Phe | Arg | Leu | Ala | Gly | 465 | 470 | 475 |
| His | Pro | Phe | Asn | Leu | Asn | Ser | Arg | Asp | Gln | Leu | Glu | Arg | Val | Leu | Phe | 485 | 490 | 495 |
| Asp | Glu | Leu | Gly | Leu | Pro | Ala | Ile | Gly | Lys | Thr | Gln | Lys | Thr | Gly | Lys | 500 | 505 | 510 |
| Arg | Ser | Thr | Ser | Ala | Ala | Val | Leu | Glu | Ala | Leu | Arg | Glu | Ala | His | Pro | 515 | 520 | 525 |
| Ile | Val | Glu | Lys | Ile | Leu | Gln | Tyr | Arg | Glu | Leu | Thr | Lys | Leu | Lys | Ser | 530 | 535 | 540 |

<210> 33
 <211> 839
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 33

```

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1          5          10          15

Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
          20          25          30

Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
          35          40          45

Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
          50          55          60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
65          70          75          80

Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
          85          90          95

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
          100          105          110

Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
          115          120          125

Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
130          135          140

Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
145          150          155          160

Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
          165          170          175

Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
          180          185          190

Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
195          200          205

Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
210          215          220

Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
225          230          235          240

Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
          245          250          255
  
```

| | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| Val | Asp | Phe | Ala | Lys | Arg | Arg | Glu | Pro | Asp | Arg | Glu | Arg | Leu | Arg | Ala | |
| | | | 260 | | | | | 265 | | | | | | 270 | | |
| Phe | Leu | Glu | Arg | Leu | Glu | Phe | Gly | Ser | Leu | Leu | His | Glu | Phe | Gly | Leu | |
| | | 275 | | | | | 280 | | | | | 285 | | | | |
| Leu | Glu | Ser | Pro | Lys | Ala | Leu | Glu | Glu | Ala | Pro | Trp | Pro | Pro | Pro | Glu | |
| | 290 | | | | | 295 | | | | | 300 | | | | | |
| Gly | Ala | Phe | Val | Gly | Phe | Val | Leu | Ser | Arg | Lys | Glu | Pro | Met | Trp | Ala | |
| 305 | | | | | 310 | | | | | 315 | | | | | 320 | |
| Asp | Leu | Leu | Ala | Leu | Ala | Ala | Ala | Arg | Gly | Gly | Arg | Val | His | Arg | Ala | |
| | | | 325 | | | | | | 330 | | | | | 335 | | |
| Pro | Glu | Pro | Tyr | Lys | Ala | Leu | Arg | Asp | Leu | Lys | Glu | Ala | Arg | Gly | Leu | |
| | | | 340 | | | | | 345 | | | | | 350 | | | |
| Leu | Ala | Lys | Asp | Leu | Ser | Val | Leu | Ala | Leu | Arg | Glu | Gly | Leu | Gly | Leu | |
| | | 355 | | | | | 360 | | | | | 365 | | | | |
| Pro | Pro | Gly | Asp | Asp | Pro | Met | Leu | Leu | Ala | Tyr | Leu | Leu | Asp | Pro | Ser | |
| | 370 | | | | | 375 | | | | | 380 | | | | | |
| Asn | Thr | Thr | Pro | Glu | Gly | Val | Ala | Arg | Arg | Tyr | Gly | Gly | Glu | Trp | Thr | |
| 385 | | | | | 390 | | | | | 395 | | | | | 400 | |
| Glu | Glu | Ala | Gly | Glu | Arg | Ala | Ala | Leu | Ser | Glu | Arg | Leu | Phe | Ala | Asn | |
| | | | | 405 | | | | | 410 | | | | | 415 | | |
| Leu | Trp | Lys | Arg | Leu | Glu | Gly | Glu | Glu | Arg | Leu | Leu | Trp | Leu | Tyr | Arg | |
| | | | 420 | | | | | 425 | | | | | 430 | | | |
| Glu | Val | Glu | Arg | Pro | Leu | Ser | Ala | Val | Leu | Ala | His | Met | Glu | Ala | Thr | |
| | | 435 | | | | | 440 | | | | | 445 | | | | |
| Gly | Val | Arg | Leu | Asp | Val | Ala | Tyr | Leu | Arg | Ala | Leu | Ser | Leu | Glu | Val | |
| | 450 | | | | | 455 | | | | | 460 | | | | | |
| Ala | Glu | Glu | Ile | Ala | Arg | Leu | Glu | Ala | Glu | Val | Phe | Arg | Leu | Ala | Gly | |
| 465 | | | | | 470 | | | | | 475 | | | | | 480 | |
| His | Pro | Phe | Asn | Leu | Asn | Ser | Arg | Asp | Gln | Leu | Glu | Arg | Val | Leu | Phe | |
| | | | 485 | | | | | | 490 | | | | | 495 | | |
| Asp | Glu | Leu | Gly | Leu | Pro | Ala | Ile | Gly | Lys | Thr | Gln | Lys | Thr | Gly | Lys | |
| | | | 500 | | | | | 505 | | | | | 510 | | | |
| Arg | Ser | Thr | Ser | Ala | Ala | Val | Leu | Glu | Ala | Leu | Arg | Glu | Ala | His | Pro | |
| | | 515 | | | | | 520 | | | | | 525 | | | | |
| Ile | Val | Glu | Lys | Ile | Leu | Gln | Tyr | Arg | Glu | Leu | Thr | Lys | Leu | Lys | Ser | |
| | 530 | | | | | 535 | | | | | 540 | | | | | |
| Thr | Tyr | Ile | Asp | Pro | Leu | Pro | Asp | Leu | Ile | His | Pro | Arg | Thr | Gly | Arg | |
| 545 | | | | | 550 | | | | | 555 | | | | | 560 | |

Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser
 565 570 575
 Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly
 580 585 590
 Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val
 595 600 605
 Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser
 610 615 620
 Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His
 625 630 635 640
 Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
 645 650 655
 Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
 660 665 670
 Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
 675 680 685
 Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
 690 695 700
 Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
 705 710 715 720
 Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
 725 730 735
 Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
 740 745 750
 Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
 755 760 765
 Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
 770 775 780
 His Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
 785 790 795 800
 Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
 805 810 815
 Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
 820 825 830
 Glu His His His His His His
 835

<210> 34
 <211> 842

<212> PRT
<213> Artificial

<220>
<223> Synthetic

<400> 34

Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
1 5 10 15
Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu
20 25 30
Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
35 40 45
Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala
50 55 60
Val Phe Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala
65 70 75 80
Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro
85 90 95
Arg Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr
100 105 110
Arg Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu
115 120 125
Ala Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala
130 135 140
Asp Arg Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His
145 150 155 160
Pro Glu Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly
165 170 175
Leu Arg Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro
180 185 190
Ser Asp Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu
195 200 205
Lys Leu Leu Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu
210 215 220
Asp Arg Val Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu
225 230 235 240
Glu Asp Leu Arg Leu Ser Leu Glu Leu Ser Arg Val Arg Thr Asp Leu
245 250 255
Pro Leu Glu Val Asp Leu Ala Gln Gly Arg Glu Pro Asp Arg Glu Gly

| | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| 260 | | | | | | | 265 | | | | | 270 | | | | | |
| Leu | Arg | Ala | Phe | Leu | Glu | Arg | Leu | Glu | Phe | Gly | Ser | Leu | Leu | His | Glu | | |
| | | 275 | | | | | 280 | | | | 285 | | | | | | |
| Phe | Gly | Leu | Leu | Glu | Ala | Pro | Ala | Pro | Leu | Glu | Glu | Ala | Pro | Trp | Pro | | |
| | | 290 | | | | | 295 | | | | 300 | | | | | | |
| Pro | Pro | Glu | Gly | Ala | Phe | Val | Gly | Phe | Val | Leu | Ser | Arg | Pro | Glu | Pro | | |
| 305 | | | | | 310 | | | | 315 | | | 320 | | | | | |
| Met | Trp | Ala | Glu | Leu | Lys | Ala | Leu | Ala | Ala | Cys | Arg | Asp | Gly | Arg | Val | | |
| | | 325 | | | | | | | 330 | | | 335 | | | | | |
| His | Arg | Ala | Ala | Asp | Pro | Leu | Ala | Gly | Leu | Lys | Asp | Leu | Lys | Glu | Val | | |
| | | 340 | | | | | | | 345 | | | 350 | | | | | |
| Arg | Gly | Leu | Leu | Ala | Lys | Asp | Leu | Ala | Val | Leu | Ala | Ser | Arg | Glu | Gly | | |
| | | 355 | | | | 360 | | | 365 | | | | | | | | |
| Leu | Asp | Leu | Val | Pro | Gly | Asp | Asp | Pro | Met | Leu | Leu | Ala | Tyr | Leu | Leu | | |
| 370 | | | | | | 375 | | | 380 | | | | | | | | |
| Asp | Pro | Ser | Asn | Thr | Thr | Pro | Glu | Gly | Val | Ala | Arg | Arg | Tyr | Gly | Gly | | |
| 385 | | | | | 390 | | | 395 | | | | | | | | 400 | |
| Glu | Trp | Thr | Glu | Asp | Ala | Ala | His | Arg | Ala | Leu | Leu | Ser | Glu | Arg | Leu | | |
| | | 405 | | | | | | | 410 | | | 415 | | | | | |
| His | Arg | Asn | Leu | Leu | Lys | Arg | Leu | Glu | Gly | Glu | Glu | Lys | Leu | Leu | Trp | | |
| | | 420 | | | | | | | 425 | | | 430 | | | | | |
| Leu | Tyr | His | Glu | Val | Glu | Lys | Pro | Leu | Ser | Arg | Val | Leu | Ala | His | Met | | |
| | | 435 | | | | 440 | | | 445 | | | | | | | | |
| Glu | Ala | Thr | Gly | Val | Arg | Arg | Asp | Val | Ala | Tyr | Leu | Gln | Ala | Leu | Ser | | |
| 450 | | | | | | 455 | | | 460 | | | | | | | | |
| Leu | Glu | Leu | Ala | Glu | Glu | Ile | Arg | Arg | Leu | Glu | Glu | Glu | Val | Phe | Arg | | |
| 465 | | | | | | 470 | | | 475 | | | | | | | 480 | |
| Leu | Ala | Gly | His | Pro | Phe | Asn | Leu | Asn | Ser | Arg | Asp | Gln | Leu | Glu | Arg | | |
| | | 485 | | | | | | | 490 | | | 495 | | | | | |
| Val | Leu | Phe | Asp | Glu | Leu | Arg | Leu | Pro | Ala | Leu | Gly | Lys | Thr | Gln | Lys | | |
| | | 500 | | | | | | | 505 | | | 510 | | | | | |
| Thr | Gly | Lys | Arg | Ser | Thr | Ser | Ala | Ala | Val | Leu | Glu | Ala | Leu | Arg | Glu | | |
| | | 515 | | | | 520 | | | 525 | | | | | | | | |
| Ala | His | Pro | Ile | Val | Glu | Lys | Ile | Leu | Gln | His | Arg | Glu | Leu | Thr | Lys | | |
| 530 | | | | | | 535 | | | 540 | | | | | | | | |
| Leu | Lys | Asn | Thr | Tyr | Val | Asp | Pro | Leu | Pro | Ser | Leu | Val | His | Pro | Arg | | |
| 545 | | | | | 550 | | | 555 | | | | | | | 560 | | |
| Thr | Gly | Arg | Leu | His | Thr | Arg | Phe | Asn | Gln | Thr | Ala | Thr | Ala | Thr | Glu | | |

<220>

<223> Synthetic

<400> 35

Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
1 5 10 15
Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu
20 25 30
Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
35 40 45
Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala
50 55 60
Val Phe Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala
65 70 75 80
Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro
85 90 95
Arg Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr
100 105 110
Arg Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu
115 120 125
Ala Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala
130 135 140
Asp Arg Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His
145 150 155 160
Pro Glu Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly
165 170 175
Leu Arg Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro
180 185 190
Ser Asp Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu
195 200 205
Lys Leu Leu Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu
210 215 220
Asp Arg Val Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu
225 230 235 240
Glu Asp Leu Arg Leu Ser Leu Glu Leu Ser Arg Val Arg Thr Asp Leu
245 250 255
Pro Leu Glu Val Asp Leu Ala Gln Gly Arg Glu Pro Asp Arg Glu Gly
260 265 270

Leu Arg Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu
 275 280 285
 Phe Gly Leu Leu Glu Ala Pro Ala Pro Leu Glu Glu Ala Pro Trp Pro
 290 295 300
 Pro Pro Glu Gly Ala Phe Val Gly Phe Val Leu Ser Arg Pro Glu Pro
 305 310 315 320
 Met Trp Ala Glu Leu Lys Ala Leu Ala Ala Cys Arg Gly Gly Arg Val
 325 330 335
 His Arg Ala Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val
 340 345 350
 Arg Gly Leu Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly
 355 360 365
 Leu Asp Leu Val Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu
 370 375 380
 Asp Pro Ser Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly
 385 390 395 400
 Glu Trp Thr Glu Asp Ala Ala His Arg Ala Leu Leu Ser Glu Arg Leu
 405 410 415
 His Arg Asn Leu Leu Lys Arg Leu Glu Gly Glu Glu Lys Leu Leu Trp
 420 425 430
 Leu Tyr His Glu Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met
 435 440 445
 Glu Ala Thr Gly Val Arg Arg Asp Val Ala Tyr Leu Gln Ala Leu Ser
 450 455 460
 Leu Glu Leu Ala Glu Glu Ile Arg Arg Leu Glu Glu Glu Val Phe Arg
 465 470 475 480
 Leu Ala Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg
 485 490 495
 Val Leu Phe Asp Glu Leu Arg Leu Pro Ala Leu Gly Lys Thr Gln Lys
 500 505 510
 Thr Gly Lys Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu
 515 520 525
 Ala His Pro Ile Val Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys
 530 535 540
 Leu Lys Asn Thr Tyr Val Asp Pro Leu Pro Ser Leu Val His Pro Arg
 545 550 555 560
 Thr Gly Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly
 565 570 575

<223> Synthetic

<400> 36

Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
1 5 10 15
Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu
20 25 30
Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
35 40 45
Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala
50 55 60
Val Phe Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala
65 70 75 80
Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro
85 90 95
Arg Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr
100 105 110
Arg Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu
115 120 125
Ala Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala
130 135 140
Asp Arg Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His
145 150 155 160
Pro Glu Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly
165 170 175
Leu Arg Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro
180 185 190
Ser Asp Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu
195 200 205
Lys Leu Leu Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu
210 215 220
Asp Arg Val Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu
225 230 235 240
Glu Asp Leu Arg Leu Ser Leu Glu Leu Ser Arg Val Arg Thr Asp Leu
245 250 255
Pro Leu Glu Val Asp Leu Ala Gln Gly Arg Glu Pro Asp Arg Glu Gly
260 265 270
Leu Arg Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu
275 280 285

| | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| 595 | | | | | 600 | | | | | 605 | | | | | | |
| Ala | Leu | Val | Ala | Leu | Asp | Tyr | Ser | Gln | Ile | Glu | Leu | Arg | Val | Leu | Ala | |
| 610 | | | | | 615 | | | | | 620 | | | | | | |
| His | Leu | Ser | Gly | Asp | Glu | Asn | Leu | Ile | Arg | Val | Phe | Gln | Glu | Gly | Lys | |
| 625 | | | | | 630 | | | | | 635 | | | | | 640 | |
| Asp | Ile | His | Thr | Gln | Thr | Ala | Ser | Trp | Met | Phe | Gly | Val | Pro | Pro | Glu | |
| 645 | | | | | 650 | | | | | 655 | | | | | | |
| Ala | Val | Asp | Pro | Leu | Met | Arg | Arg | Ala | Ala | Lys | Thr | Val | Asn | Phe | Gly | |
| 660 | | | | | 665 | | | | | 670 | | | | | | |
| Val | Leu | Tyr | Gly | Met | Ser | Ala | His | Arg | Leu | Ser | Gln | Glu | Leu | Ala | Ile | |
| 675 | | | | | 680 | | | | | 685 | | | | | | |
| Pro | Tyr | Glu | Glu | Ala | Val | Ala | Phe | Ile | Glu | Arg | Tyr | Phe | Gln | Ser | Phe | |
| 690 | | | | | 695 | | | | | 700 | | | | | | |
| Pro | Lys | Val | Arg | Ala | Trp | Ile | Glu | Lys | Thr | Leu | Glu | Glu | Gly | Arg | Lys | |
| 705 | | | | | 710 | | | | | 715 | | | | | 720 | |
| Arg | Gly | Tyr | Val | Glu | Thr | Leu | Phe | Gly | Arg | Arg | Arg | Tyr | Val | Pro | Asp | |
| 725 | | | | | 730 | | | | | 735 | | | | | | |
| Leu | Asn | Ala | Arg | Val | Lys | Ser | Val | Arg | Glu | Ala | Ala | Glu | Arg | Met | Ala | |
| 740 | | | | | 745 | | | | | 750 | | | | | | |
| Phe | Asn | Met | Pro | Val | Gln | Gly | Thr | Ala | Ala | Asp | Leu | Met | Lys | Leu | Ala | |
| 755 | | | | | 760 | | | | | 765 | | | | | | |
| Met | Val | Lys | Leu | Phe | Pro | Arg | Leu | Arg | Glu | Met | Gly | Ala | Arg | Met | Leu | |
| 770 | | | | | 775 | | | | | 780 | | | | | | |
| Leu | Gln | Val | Ala | Asn | Glu | Leu | Leu | Leu | Glu | Ala | Pro | Gln | Ala | Arg | Ala | |
| 785 | | | | | 790 | | | | | 795 | | | | | 800 | |
| Glu | Glu | Val | Ala | Ala | Leu | Ala | Lys | Glu | Ala | Met | Glu | Lys | Ala | Tyr | Pro | |
| 805 | | | | | 810 | | | | | 815 | | | | | | |
| Leu | Ala | Val | Pro | Leu | Glu | Val | Glu | Val | Gly | Met | Gly | Glu | Asp | Trp | Leu | |
| 820 | | | | | 825 | | | | | 830 | | | | | | |
| Ser | Ala | Lys | Gly | His | His | His | His | His | His | | | | | | | |
| 835 | | | | | 840 | | | | | | | | | | | |

```
<210> 38
<211> 839
<212> PRT
<213> Artificial
```

<220>
<223> Synthetic

<400> 38

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Asn | Ser | Gly | Met | Leu | Pro | Leu | Phe | Glu | Pro | Lys | Gly | Arg | Val | Leu | 1 | 5 | 10 | 15 |
| Leu | Val | Asp | Gly | His | His | Leu | Ala | Tyr | Arg | Thr | Phe | His | Ala | Leu | Lys | 20 | 25 | 30 | |
| Gly | Leu | Thr | Thr | Ser | Arg | Gly | Glu | Pro | Val | Gln | Ala | Val | Tyr | Gly | Phe | 35 | 40 | 45 | |
| Ala | Lys | Ser | Leu | Leu | Lys | Ala | Leu | Lys | Glu | Asp | Gly | Asp | Ala | Val | Ile | 50 | 55 | 60 | |
| Val | Val | Phe | Asp | Ala | Lys | Ala | Pro | Ser | Phe | Arg | His | Glu | Ala | Tyr | Gly | 65 | 70 | 75 | 80 |
| Gly | Tyr | Lys | Ala | Gly | Arg | Ala | Pro | Thr | Pro | Glu | Asp | Phe | Pro | Arg | Gln | 85 | 90 | 95 | |
| Leu | Ala | Leu | Ile | Lys | Glu | Leu | Val | Asp | Leu | Leu | Gly | Leu | Ala | Arg | Leu | 100 | 105 | 110 | |
| Glu | Val | Pro | Gly | Tyr | Glu | Ala | Asp | Asp | Val | Leu | Ala | Ser | Leu | Ala | Lys | 115 | 120 | 125 | |
| Lys | Ala | Glu | Lys | Glu | Gly | Tyr | Glu | Val | Arg | Ile | Leu | Thr | Ala | Asp | Lys | 130 | 135 | 140 | |
| Asp | Leu | Tyr | Gln | Leu | Leu | Ser | Asp | Arg | Ile | His | Val | Leu | His | Pro | Glu | 145 | 150 | 155 | 160 |
| Gly | Tyr | Leu | Ile | Thr | Pro | Ala | Trp | Leu | Trp | Glu | Lys | Tyr | Gly | Leu | Arg | 165 | 170 | 175 | |
| Pro | Asp | Gln | Trp | Ala | Asp | Tyr | Arg | Ala | Leu | Thr | Gly | Asp | Glu | Ser | Asp | 180 | 185 | 190 | |
| Asn | Leu | Pro | Gly | Val | Lys | Gly | Ile | Gly | Glu | Lys | Thr | Ala | Arg | Lys | Leu | 195 | 200 | 205 | |
| Leu | Glu | Glu | Trp | Gly | Ser | Leu | Glu | Ala | Leu | Leu | Lys | Asn | Leu | Asp | Arg | 210 | 215 | 220 | |
| Leu | Lys | Pro | Ala | Ile | Arg | Glu | Lys | Ile | Leu | Ala | His | Met | Asp | Asp | Leu | 225 | 230 | 235 | 240 |
| Lys | Leu | Ser | Trp | Asp | Leu | Ala | Lys | Val | Arg | Thr | Asp | Leu | Pro | Leu | Glu | 245 | 250 | 255 | |
| Val | Asp | Phe | Ala | Lys | Arg | Arg | Glu | Pro | Asp | Arg | Glu | Arg | Leu | Arg | Ala | 260 | 265 | 270 | |
| Phe | Leu | Glu | Arg | Leu | Glu | Phe | Gly | Ser | Leu | Leu | His | Glu | Phe | Gly | Leu | 275 | 280 | 285 | |
| Leu | Glu | Ser | Pro | Lys | Ala | Leu | Glu | Glu | Ala | Pro | Trp | Pro | Pro | Pro | Glu | 290 | 295 | 300 | |

| | | | | | | | | | | | | | | | |
|-------|------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ala | Leu | Asp | Tyr | Ser | Gln | Ile | Glu | Leu | Arg | Val | Leu | Ala | His | Leu | Ser |
| 610 | | | | | | 615 | | | | | 620 | | | | |
| Gly | Asp | Glu | Asn | Leu | Ile | Arg | Val | Phe | Gln | Glu | Gly | Arg | Asp | Ile | His |
| 625 | | | | | 630 | | | | | 635 | | | | 640 | |
| Thr | Glu | Thr | Ala | Ser | Trp | Met | Phe | Gly | Val | Pro | Arg | Glu | Ala | Val | Asp |
| | | | | 645 | | | | | 650 | | | | | 655 | |
| Pro | Leu | Met | Arg | Arg | Ala | Ala | Lys | Thr | Ile | Asn | Phe | Gly | Val | Leu | Tyr |
| | | | 660 | | | | | 665 | | | | | 670 | | |
| Gly | Met | Ser | Ala | His | Arg | Leu | Ser | Gln | Glu | Leu | Ala | Ile | Pro | Tyr | Glu |
| | | 675 | | | | | 680 | | | | | 685 | | | |
| Glu | Ala | Gln | Ala | Phe | Ile | Glu | Arg | Tyr | Phe | Gln | Ser | Phe | Pro | Lys | Val |
| | 690 | | | | | 695 | | | | | 700 | | | | |
| Arg | Ala | Trp | Ile | Glu | Lys | Thr | Leu | Glu | Glu | Gly | Arg | Arg | Arg | Gly | Tyr |
| 705 | | | | | 710 | | | | | 715 | | | | | 720 |
| Val | Glu | Thr | Leu | Phe | Gly | Arg | Arg | Arg | Tyr | Val | Pro | Asp | Leu | Glu | Ala |
| | | | | 725 | | | | | 730 | | | | | 735 | |
| Arg | Val | Lys | Ser | Val | Arg | Glu | Ala | Ala | Glu | Arg | Met | Ala | Phe | Asn | Met |
| | | | 740 | | | | | 745 | | | | | 750 | | |
| Pro | Val | Gln | Gly | Thr | Ala | Ala | Asp | Leu | Met | Lys | Leu | Ala | Met | Val | Lys |
| | | 755 | | | | | 760 | | | | | 765 | | | |
| Leu | Phe | Pro | Arg | Leu | Glu | Glu | Met | Gly | Ala | Arg | Met | Leu | Leu | Gln | Val |
| | 770 | | | | | 775 | | | | | 780 | | | | |
| Ala | Asn | Glu | Leu | Val | Leu | Glu | Ala | Pro | Lys | Glu | Arg | Ala | Glu | Ala | Val |
| 785 | | | | | 790 | | | | | 795 | | | | | 800 |
| Ala | Arg | Leu | Ala | Lys | Glu | Val | Met | Glu | Gly | Val | Tyr | Pro | Leu | Ala | Val |
| | | | | 805 | | | | | 810 | | | | | 815 | |
| Pro | Leu | Glu | Val | Glu | Val | Gly | Ile | Gly | Glu | Asp | Trp | Leu | Ser | Ala | Lys |
| | | 820 | | | | | | 825 | | | | | 830 | | |
| Glu | His | His | His | His | His | His | | | | | | | | | |
| | | 835 | | | | | | | | | | | | | |
| <210> | 39 | | | | | | | | | | | | | | |
| <211> | 839 | | | | | | | | | | | | | | |
| <212> | PRT | | | | | | | | | | | | | | |
| <213> | Artificial | | | | | | | | | | | | | | |
| <220> | | | | | | | | | | | | | | | |
| <223> | Synthetic | | | | | | | | | | | | | | |
| <400> | 39 | | | | | | | | | | | | | | |
| Met | Asn | Ser | Gly | Met | Leu | Pro | Leu | Phe | Glu | Pro | Lys | Gly | Arg | Val | Leu |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |

Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala
 325 330 335
 Pro Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu
 340 345 350
 Leu Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu
 355 360 365
 Pro Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser
 370 375 380
 Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr
 385 390 395 400
 Glu Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn
 405 410 415
 Leu Leu Lys Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg
 420 425 430
 Glu Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr
 435 440 445
 Gly Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val
 450 455 460
 Ala Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly
 465 470 475 480
 His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe
 485 490 495
 Asp Glu Leu Gly Leu Pro Ala Ile Gly Lys Thr Gln Lys Thr Gly Lys
 500 505 510
 Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro
 515 520 525
 Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser
 530 535 540
 Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg
 545 550 555 560
 Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser
 565 570 575
 Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly
 580 585 590
 Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val
 595 600 605
 Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser
 610 615 620

| 20 | | | | | 25 | | | | | 30 | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| Gly | Leu | Thr | Thr | Ser | Arg | Gly | Glu | Pro | Val | Gln | Ala | Val | Tyr | Gly | Phe | |
| 35 | | | | | 40 | | | | | 45 | | | | | | |
| Ala | Lys | Ser | Leu | Leu | Lys | Ala | Leu | Lys | Glu | Asp | Gly | Asp | Ala | Val | Ile | |
| 50 | | | | | 55 | | | | | 60 | | | | | | |
| Val | Val | Phe | Asp | Ala | Lys | Ala | Pro | Ser | Phe | Arg | His | Glu | Ala | Tyr | Gly | |
| 65 | | | | | 70 | | | | | 75 | | | | | 80 | |
| Gly | Tyr | Lys | Ala | Gly | Arg | Ala | Pro | Thr | Pro | Glu | Asp | Phe | Pro | Arg | Gln | |
| 85 | | | | | 90 | | | | | 95 | | | | | | |
| Leu | Ala | Leu | Ile | Lys | Glu | Leu | Val | Asp | Leu | Leu | Gly | Leu | Ala | Arg | Leu | |
| 100 | | | | | 105 | | | | | 110 | | | | | | |
| Glu | Val | Pro | Gly | Tyr | Glu | Ala | Asp | Asp | Val | Leu | Ala | Ser | Leu | Ala | Lys | |
| 115 | | | | | 120 | | | | | 125 | | | | | | |
| Lys | Ala | Glu | Lys | Glu | Gly | Tyr | Glu | Val | Arg | Ile | Leu | Thr | Ala | Asp | Lys | |
| 130 | | | | | 135 | | | | | 140 | | | | | | |
| Asp | Leu | Tyr | Gln | Leu | Leu | Ser | Asp | Arg | Ile | His | Val | Leu | His | Pro | Glu | |
| 145 | | | | | 150 | | | | | 155 | | | | | 160 | |
| Gly | Tyr | Leu | Ile | Thr | Pro | Ala | Trp | Leu | Trp | Glu | Lys | Tyr | Gly | Leu | Arg | |
| 165 | | | | | 170 | | | | | 175 | | | | | | |
| Pro | Asp | Gln | Trp | Ala | Asp | Tyr | Arg | Ala | Leu | Thr | Gly | Asp | Glu | Ser | Asp | |
| 180 | | | | | 185 | | | | | 190 | | | | | | |
| Asn | Leu | Pro | Gly | Val | Lys | Gly | Ile | Gly | Glu | Lys | Thr | Ala | Arg | Lys | Leu | |
| 195 | | | | | 200 | | | | | 205 | | | | | | |
| Leu | Glu | Glu | Trp | Gly | Ser | Leu | Glu | Ala | Leu | Leu | Lys | Asn | Leu | Asp | Arg | |
| 210 | | | | | 215 | | | | | 220 | | | | | | |
| Leu | Lys | Pro | Ala | Ile | Arg | Glu | Lys | Ile | Leu | Ala | His | Met | Asp | Asp | Leu | |
| 225 | | | | | 230 | | | | | 235 | | | | | 240 | |
| Lys | Leu | Ser | Trp | Asp | Leu | Ala | Lys | Val | Arg | Thr | Asp | Leu | Pro | Leu | Glu | |
| 245 | | | | | 250 | | | | | 255 | | | | | | |
| Val | Asp | Phe | Ala | Lys | Arg | Arg | Glu | Pro | Asp | Arg | Glu | Arg | Leu | Arg | Ala | |
| 260 | | | | | 265 | | | | | 270 | | | | | | |
| Phe | Leu | Glu | Arg | Leu | Glu | Phe | Gly | Ser | Leu | Leu | His | Glu | Phe | Gly | Leu | |
| 275 | | | | | 280 | | | | | 285 | | | | | | |
| Leu | Glu | Ser | Pro | Lys | Ala | Leu | Glu | Glu | Ala | Pro | Trp | Pro | Pro | Pro | Glu | |
| 290 | | | | | 295 | | | | | 300 | | | | | | |
| Gly | Ala | Phe | Val | Gly | Phe | Val | Leu | Ser | Arg | Lys | Glu | Pro | Met | Trp | Ala | |
| 305 | | | | | 310 | | | | | 315 | | | | | 320 | |
| Asp | Leu | Leu | Ala | Leu | Ala | Ala | Ala | Arg | Gly | Gly | Arg | Val | His | Arg | Ala | |

| | | | | | |
|---|-----|--|-----|--|-----|
| | 325 | | 330 | | 335 |
| Pro Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu | 340 | | 345 | | 350 |
| Leu Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu | 355 | | 360 | | 365 |
| Pro Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser | 370 | | 375 | | 380 |
| Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr | 385 | | 390 | | 395 |
| Glu Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn | 405 | | 410 | | 415 |
| Leu Leu Lys Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg | 420 | | 425 | | 430 |
| Glu Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr | 435 | | 440 | | 445 |
| Gly Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val | 450 | | 455 | | 460 |
| Ala Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly | 465 | | 470 | | 475 |
| His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe | 485 | | 490 | | 495 |
| Asp Glu Leu Gly Leu Pro Ala Ile Gly Lys Thr Gln Lys Thr Gly Lys | 500 | | 505 | | 510 |
| Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro | 515 | | 520 | | 525 |
| Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser | 530 | | 535 | | 540 |
| Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg | 545 | | 550 | | 555 |
| Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser | 565 | | 570 | | 575 |
| Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Ala Thr Pro Leu Gly | 580 | | 585 | | 590 |
| Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val | 595 | | 600 | | 605 |
| Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser | 610 | | 615 | | 620 |
| Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His | | | | | |

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Gly | Leu | Thr | Thr | Ser | Arg | Gly | Glu | Pro | Val | Gln | Ala | Val | Tyr | Gly | Phe | 35 | 40 | 45 | |
| Ala | Lys | Ser | Leu | Leu | Lys | Ala | Leu | Lys | Glu | Asp | Gly | Asp | Ala | Val | Ile | 50 | 55 | 60 | |
| Val | Val | Phe | Asp | Ala | Lys | Ala | Pro | Ser | Phe | Arg | His | Glu | Ala | Tyr | Gly | 65 | 70 | 75 | 80 |
| Gly | Tyr | Lys | Ala | Gly | Arg | Ala | Pro | Thr | Pro | Glu | Asp | Phe | Pro | Arg | Gln | 85 | 90 | 95 | |
| Leu | Ala | Leu | Ile | Lys | Glu | Leu | Val | Asp | Leu | Leu | Gly | Leu | Ala | Arg | Leu | 100 | 105 | 110 | |
| Glu | Val | Pro | Gly | Tyr | Glu | Ala | Asp | Asp | Val | Leu | Ala | Ser | Leu | Ala | Lys | 115 | 120 | 125 | |
| Lys | Ala | Glu | Lys | Glu | Gly | Tyr | Glu | Val | Arg | Ile | Leu | Thr | Ala | Asp | Lys | 130 | 135 | 140 | |
| Asp | Leu | Tyr | Gln | Leu | Leu | Ser | Asp | Arg | Ile | His | Val | Leu | His | Pro | Glu | 145 | 150 | 155 | 160 |
| Gly | Tyr | Leu | Ile | Thr | Pro | Ala | Trp | Leu | Trp | Glu | Lys | Tyr | Gly | Leu | Arg | 165 | 170 | 175 | |
| Pro | Asp | Gln | Trp | Ala | Asp | Tyr | Arg | Ala | Leu | Thr | Gly | Asp | Glu | Ser | Asp | 180 | 185 | 190 | |
| Asn | Leu | Pro | Gly | Val | Lys | Gly | Ile | Gly | Glu | Lys | Thr | Ala | Arg | Lys | Leu | 195 | 200 | 205 | |
| Leu | Glu | Glu | Trp | Gly | Ser | Leu | Glu | Ala | Leu | Leu | Lys | Asn | Leu | Asp | Arg | 210 | 215 | 220 | |
| Leu | Lys | Pro | Ala | Ile | Arg | Glu | Lys | Ile | Leu | Ala | His | Met | Asp | Asp | Leu | 225 | 230 | 235 | 240 |
| Lys | Leu | Ser | Trp | Asp | Leu | Ala | Lys | Val | Arg | Thr | Asp | Leu | Pro | Leu | Glu | 245 | 250 | 255 | |
| Val | Asp | Phe | Ala | Lys | Arg | Arg | Glu | Pro | Asp | Arg | Glu | Arg | Leu | Arg | Ala | 260 | 265 | 270 | |
| Phe | Leu | Glu | Arg | Leu | Glu | Phe | Gly | Ser | Leu | Leu | His | Glu | Phe | Gly | Leu | 275 | 280 | 285 | |
| Leu | Glu | Ser | Pro | Lys | Ala | Leu | Glu | Glu | Ala | Pro | Trp | Pro | Pro | Pro | Glu | 290 | 295 | 300 | |
| Gly | Ala | Phe | Val | Gly | Phe | Val | Leu | Ser | Arg | Lys | Glu | Pro | Met | Trp | Ala | 305 | 310 | 315 | 320 |
| Asp | Leu | Leu | Ala | Leu | Ala | Ala | Ala | Arg | Gly | Gly | Arg | Val | His | Arg | Ala | 325 | 330 | 335 | |

Pro Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu
340 345 350

Leu Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu
355 360 365

Pro Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser
370 375 380

Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr
385 390 395 400

Glu Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn
405 410 415

Leu Leu Lys Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg
420 425 430

Glu Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr
435 440 445

Gly Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val
450 455 460

Ala Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly
465 470 475 480

His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe
485 490 495

Asp Glu Leu Gly Leu Pro Ala Ile Lys Lys Thr Gln Lys Thr Gly Lys
500 505 510

Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro
515 520 525

Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser
530 535 540

Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg
545 550 555 560

Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser
565 570 575

Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly
580 585 590

Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val
595 600 605

Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser
610 615 620

Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His
625 630 635 640

Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ala | Lys | Ser | Leu | Leu | Lys | Ala | Leu | Lys | Glu | Asp | Gly | Asp | Ala | Val | Ile | 50 | 55 | 60 | |
| Val | Val | Phe | Asp | Ala | Lys | Ala | Pro | Ser | Phe | Arg | His | Glu | Ala | Tyr | Glu | 65 | 70 | 75 | 80 |
| Gly | Tyr | Lys | Ala | Gly | Arg | Ala | Pro | Thr | Pro | Glu | Asp | Phe | Pro | Arg | Gln | 85 | 90 | 95 | |
| Leu | Ala | Leu | Ile | Lys | Glu | Leu | Val | Asp | Leu | Leu | Gly | Leu | Ala | Arg | Leu | 100 | 105 | 110 | |
| Glu | Val | Pro | Gly | Tyr | Glu | Ala | Asp | Asp | Val | Leu | Ala | Ser | Leu | Ala | Lys | 115 | 120 | 125 | |
| Lys | Ala | Glu | Lys | Glu | Gly | Tyr | Glu | Val | Arg | Ile | Leu | Thr | Ala | Asp | Lys | 130 | 135 | 140 | |
| Asp | Leu | Tyr | Gln | Leu | Leu | Ser | Asp | Arg | Ile | His | Val | Leu | His | Pro | Glu | 145 | 150 | 155 | 160 |
| Gly | Tyr | Leu | Ile | Thr | Pro | Ala | Trp | Leu | Trp | Glu | Lys | Tyr | Gly | Leu | Arg | 165 | 170 | 175 | |
| Pro | Asp | Gln | Trp | Ala | Asp | Tyr | Arg | Ala | Leu | Thr | Gly | Asp | Glu | Ser | Asp | 180 | 185 | 190 | |
| Asn | Leu | Pro | Gly | Val | Lys | Gly | Ile | Gly | Glu | Lys | Thr | Ala | Arg | Lys | Leu | 195 | 200 | 205 | |
| Leu | Glu | Glu | Trp | Gly | Ser | Leu | Glu | Ala | Leu | Leu | Lys | Asn | Leu | Asp | Arg | 210 | 215 | 220 | |
| Leu | Lys | Pro | Ala | Ile | Arg | Glu | Lys | Ile | Leu | Ala | His | Met | Asp | Asp | Leu | 225 | 230 | 235 | 240 |
| Lys | Leu | Ser | Trp | Asp | Leu | Ala | Lys | Val | Arg | Thr | Asp | Leu | Pro | Leu | Glu | 245 | 250 | 255 | |
| Val | Asp | Phe | Ala | Lys | Arg | Arg | Glu | Pro | Asp | Arg | Glu | Arg | Leu | Arg | Ala | 260 | 265 | 270 | |
| Phe | Leu | Glu | Arg | Leu | Glu | Phe | Gly | Ser | Leu | Leu | His | Glu | Phe | Gly | Leu | 275 | 280 | 285 | |
| Leu | Glu | Ser | Pro | Lys | Ala | Leu | Glu | Glu | Ala | Pro | Trp | Pro | Pro | Pro | Glu | 290 | 295 | 300 | |
| Gly | Ala | Phe | Val | Gly | Phe | Val | Leu | Ser | Arg | Lys | Glu | Pro | Met | Trp | Ala | 305 | 310 | 315 | 320 |
| Asp | Leu | Leu | Ala | Leu | Ala | Ala | Ala | Arg | Gly | Gly | Arg | Val | His | Arg | Ala | 325 | 330 | 335 | |
| Pro | Glu | Pro | Tyr | Lys | Ala | Leu | Arg | Asp | Leu | Lys | Glu | Ala | Arg | Gly | Leu | 340 | 345 | 350 | |

Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
660 665 670

Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
675 680 685

Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
690 695 700

Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
705 710 715 720

Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
725 730 735

Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
740 745 750

Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
755 760 765

Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
770 775 780

Ala Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
785 790 795 800

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Glu His His His His His His
835

<210> 43
<211> 839
<212> PRT
<213> Artificial

<220>
<223> Synthetic

<400> 43

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1 5 10 15

Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
20 25 30

Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
35 40 45

Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
50 55 60

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Pro | Pro | Gly | Asp | Asp | Pro | Met | Leu | Leu | Ala | Tyr | Leu | Leu | Asp | Pro | Ser | 370 | 375 | 380 | |
| Asn | Thr | Thr | Pro | Glu | Gly | Val | Ala | Arg | Arg | Tyr | Gly | Gly | Glu | Trp | Thr | 385 | 390 | 395 | 400 |
| Glu | Glu | Ala | Gly | Glu | Arg | Ala | Ala | Leu | Ser | Glu | Arg | Leu | Phe | Ala | Asn | 405 | 410 | 415 | |
| Leu | Leu | Lys | Arg | Leu | Glu | Gly | Glu | Glu | Arg | Leu | Leu | Trp | Leu | Tyr | Arg | 420 | 425 | 430 | |
| Glu | Val | Glu | Arg | Pro | Leu | Ser | Ala | Val | Leu | Ala | His | Met | Glu | Ala | Thr | 435 | 440 | 445 | |
| Gly | Val | Arg | Leu | Asp | Val | Ala | Tyr | Leu | Arg | Ala | Leu | Ser | Leu | Glu | Val | 450 | 455 | 460 | |
| Ala | Glu | Glu | Ile | Ala | Arg | Leu | Glu | Ala | Glu | Val | Phe | Arg | Leu | Ala | Gly | 465 | 470 | 475 | 480 |
| His | Pro | Phe | Asn | Leu | Asn | Ser | Arg | Asp | Gln | Leu | Glu | Arg | Val | Leu | Phe | 485 | 490 | 495 | |
| Asp | Glu | Leu | Gly | Leu | Pro | Ala | Ile | Gly | Lys | Thr | Gln | Lys | Thr | Gly | Lys | 500 | 505 | 510 | |
| Arg | Ser | Thr | Ser | Ala | Ala | Val | Leu | Glu | Ala | Leu | Arg | Glu | Ala | His | Pro | 515 | 520 | 525 | |
| Ile | Val | Glu | Lys | Ile | Leu | Gln | Tyr | Arg | Glu | Leu | Thr | Lys | Leu | Lys | Ser | 530 | 535 | 540 | |
| Thr | Tyr | Ile | Asp | Pro | Leu | Pro | Asp | Leu | Ile | His | Pro | Arg | Thr | Gly | Arg | 545 | 550 | 555 | 560 |
| Leu | His | Thr | Arg | Phe | Asn | Gln | Thr | Ala | Thr | Ala | Thr | Gly | Arg | Leu | Ser | 565 | 570 | 575 | |
| Ser | Ser | Asp | Pro | Asn | Leu | Gln | Asn | Ile | Pro | Val | Arg | Thr | Pro | Leu | Gly | 580 | 585 | 590 | |
| Gln | Arg | Ile | Arg | Arg | Ala | Phe | Ile | Ala | Glu | Glu | Gly | Trp | Leu | Leu | Val | 595 | 600 | 605 | |
| Ala | Leu | Asp | Tyr | Ser | Gln | Ile | Glu | Leu | Arg | Val | Leu | Ala | His | Leu | Ser | 610 | 615 | 620 | |
| Gly | Asp | Glu | Asn | Leu | Ile | Arg | Val | Phe | Gln | Glu | Gly | Arg | Asp | Ile | His | 625 | 630 | 635 | 640 |
| Thr | Glu | Thr | Ala | Ser | Trp | Met | Phe | Gly | Val | Pro | Arg | Glu | Ala | Val | Asp | 645 | 650 | 655 | |
| Pro | Leu | Met | Arg | Arg | Ala | Ala | Lys | Thr | Ile | Asn | Phe | Gly | Val | Leu | Tyr | 660 | 665 | 670 | |

Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
675 680 685

Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
690 695 700

Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
705 710 715 720

Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
725 730 735

Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
740 745 750

Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
755 760 765

Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
770 775 780

Ala Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
785 790 795 800

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Glu His His His His His His
835

<210> 44
<211> 839
<212> PRT
<213> Artificial

<220>
<223> Synthetic

<400> 44

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1 5 10 15

Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
20 25 30

Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
35 40 45

Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
50 55 60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly

| | | |
|---|---|-----|
| 675 | 680 | 685 |
| Glu Ala Gln Ala Phe Ile | Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val | |
| 690 | 695 | 700 |
| Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr | | |
| 705 | 710 | 715 |
| Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala | | |
| 725 | 730 | 735 |
| Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met | | |
| 740 | 745 | 750 |
| Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys | | |
| 755 | 760 | 765 |
| Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val | | |
| 770 | 775 | 780 |
| Ala Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val | | |
| 785 | 790 | 795 |
| Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val | | |
| 805 | 810 | 815 |
| Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys | | |
| 820 | 825 | 830 |
| Glu His His His His His His | | |
| 835 | | |
| <210> 45 | | |
| <211> 839 | | |
| <212> PRT | | |
| <213> Artificial | | |
| <220> | | |
| <223> Synthetic | | |
| <400> 45 | | |
| Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu | | |
| 1 | 5 | 10 |
| Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys | | |
| 20 | 25 | 30 |
| Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe | | |
| 35 | 40 | 45 |
| Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile | | |
| 50 | 55 | 60 |
| Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly | | |
| 65 | 70 | 75 |
| | | 80 |

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
 100 105 110
 Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
 115 120 125
 Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
 130 135 140
 Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
 145 150 155 160
 Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
 165 170 175
 Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
 180 185 190
 Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
 195 200 205
 Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
 210 215 220
 Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
 225 230 235 240
 Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
 245 250 255
 Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
 260 265 270
 Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
 275 280 285
 Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu
 290 295 300
 Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala
 305 310 315 320
 Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala
 325 330 335
 Pro Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu
 340 345 350
 Leu Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu
 355 360 365
 Pro Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser
 370 375 380
 Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr
 385 390 395 400

Glu Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn
 405 410 415
 Leu Leu Lys Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg
 420 425 430
 Glu Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr
 435 440 445
 Gly Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val
 450 455 460
 Ala Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly
 465 470 475 480
 His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe
 485 490 495
 Asp Glu Leu Arg Leu Pro Lys Ile Asn Lys Thr Lys Lys Thr Gly Lys
 500 505 510
 Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro
 515 520 525
 Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser
 530 535 540
 Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg
 545 550 555 560
 Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser
 565 570 575
 Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly
 580 585 590
 Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val
 595 600 605
 Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser
 610 615 620
 Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His
 625 630 635 640
 Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
 645 650 655
 Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
 660 665 670
 Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
 675 680 685
 Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
 690 695 700

| | | |
|---|---------------------------------|-----|
| 100 | 105 | 110 |
| Glu Val Pro Gly Tyr Glu Ala Asp | Asp Val Leu Ala Ser Leu Ala Lys | |
| 115 | 120 | 125 |
| Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys | | |
| 130 | 135 | 140 |
| Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu | | |
| 145 | 150 | 155 |
| Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg | | |
| 165 | 170 | 175 |
| Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp | | |
| 180 | 185 | 190 |
| Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu | | |
| 195 | 200 | 205 |
| Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg | | |
| 210 | 215 | 220 |
| Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu | | |
| 225 | 230 | 235 |
| Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu | | |
| 245 | 250 | 255 |
| Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala | | |
| 260 | 265 | 270 |
| Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu | | |
| 275 | 280 | 285 |
| Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu | | |
| 290 | 295 | 300 |
| Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala | | |
| 305 | 310 | 315 |
| Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala | | |
| 325 | 330 | 335 |
| Pro Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu | | |
| 340 | 345 | 350 |
| Leu Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu | | |
| 355 | 360 | 365 |
| Pro Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser | | |
| 370 | 375 | 380 |
| Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr | | |
| 385 | 390 | 395 |
| Glu Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn | | |

| | | | | | | |
|---|--|-----|--|-----|--|-----|
| 705 | | 710 | | 715 | | 720 |
| Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala | | | | | | |
| | | 725 | | 730 | | 735 |
| Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met | | | | | | |
| | | 740 | | 745 | | 750 |
| Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys | | | | | | |
| | | 755 | | 760 | | 765 |
| Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val | | | | | | |
| | | 770 | | 775 | | 780 |
| Ala Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val | | | | | | |
| | | 785 | | 790 | | 795 |
| Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val | | | | | | |
| | | 805 | | 810 | | 815 |
| Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys | | | | | | |
| | | 820 | | 825 | | 830 |
| Glu His His His His His His | | | | | | |
| | | 835 | | | | |
| <210> 48 | | | | | | |
| <211> 839 | | | | | | |
| <212> PRT | | | | | | |
| <213> Artificial | | | | | | |
| <220> | | | | | | |
| <223> Synthetic | | | | | | |
| <400> 48 | | | | | | |
| Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu | | | | | | |
| 1 | | 5 | | 10 | | 15 |
| Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys | | | | | | |
| | | 20 | | 25 | | 30 |
| Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe | | | | | | |
| | | 35 | | 40 | | 45 |
| Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile | | | | | | |
| | | 50 | | 55 | | 60 |
| Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly | | | | | | |
| | | 65 | | 70 | | 75 |
| Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln | | | | | | |
| | | 85 | | 90 | | 95 |
| Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu | | | | | | |
| | | 100 | | 105 | | 110 |

Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
725 730 735

Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
740 745 750

Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
755 760 765

Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
770 775 780

Ala Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
785 790 795 800

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Glu His His His His His His
835

<210> 49
<211> 839
<212> PRT
<213> Artificial

<220>
<223> Synthetic

<400> 49

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1 5 10 15

Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
20 25 30

Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
35 40 45

Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
50 55 60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
65 70 75 80

Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
85 90 95

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
100 105 110

Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
115 120 125

Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
 130 135 140
 Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
 145 150 155 160
 Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
 165 170 175
 Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
 180 185 190
 Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
 195 200 205
 Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
 210 215 220
 Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
 225 230 235 240
 Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
 245 250 255
 Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
 260 265 270
 Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
 275 280 285
 Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu
 290 295 300
 Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala
 305 310 315 320
 Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala
 325 330 335
 Pro Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu
 340 345 350
 Leu Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu
 355 360 365
 Pro Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser
 370 375 380
 Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr
 385 390 395 400
 Glu Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn
 405 410 415
 Leu Leu Lys Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg
 420 425 430

Glu Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr
435 440 445

Gly Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val
450 455 460

Ala Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly
465 470 475 480

His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe
485 490 495

Asp Glu Leu Arg Leu Pro Lys Leu Lys Lys Thr Lys Lys Thr Gly Lys
500 505 510

Arg Ser Thr Ser Ala Ala Leu Leu Glu Ala Leu Arg Glu Ala His Pro
515 520 525

Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser
530 535 540

Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg
545 550 555 560

Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser
565 570 575

Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly
580 585 590

Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val
595 600 605

Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser
610 615 620

Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His
625 630 635 640

Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
645 650 655

Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
660 665 670

Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
675 680 685

Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
690 695 700

Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
705 710 715 720

Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
725 730 735

| | | |
|---|-----|-------------|
| 130 | 135 | 140 |
| Asp Arg Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His | | |
| 145 | 150 | 155 160 |
| Pro Glu Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly | | |
| | 165 | 170 175 |
| Leu Arg Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro | | |
| | 180 | 185 190 |
| Ser Asp Asn Leu Arg Gly Val Arg Gly Ile Gly Glu Lys Thr Ala Leu | | |
| | 195 | 200 205 |
| Lys Leu Leu Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu | | |
| | 210 | 215 220 |
| Asp Arg Val Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu | | |
| | 225 | 230 235 240 |
| Glu Asp Leu Arg Leu Ser Leu Glu Leu Ser Arg Val Arg Thr Asp Leu | | |
| | 245 | 250 255 |
| Pro Leu Glu Val Asp Leu Ala Gln Gly Arg Glu Pro Asp Arg Glu Gly | | |
| | 260 | 265 270 |
| Leu Arg Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu | | |
| | 275 | 280 285 |
| Phe Gly Leu Leu Glu Ala Pro Ala Pro Leu Glu Glu Ala Pro Trp Pro | | |
| | 290 | 295 300 |
| Pro Pro Glu Gly Ala Phe Val Gly Phe Val Leu Ser Arg Pro Glu Pro | | |
| | 305 | 310 315 320 |
| Met Trp Ala Glu Leu Lys Ala Leu Ala Ala Cys Arg Gly Gly Arg Val | | |
| | 325 | 330 335 |
| His Arg Ala Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val | | |
| | 340 | 345 350 |
| Arg Gly Leu Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly | | |
| | 355 | 360 365 |
| Leu Asp Leu Val Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu | | |
| | 370 | 375 380 |
| Asp Pro Ser Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly | | |
| | 385 | 390 395 400 |
| Glu Trp Thr Glu Asp Ala Ala His Arg Ala Leu Leu Ser Glu Arg Leu | | |
| | 405 | 410 415 |
| His Arg Asn Leu Leu Lys Arg Leu Glu Gly Glu Glu Lys Leu Leu Trp | | |
| | 420 | 425 430 |
| Leu Tyr His Glu Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met | | |

| | | |
|---|-----|-----|
| 740 | 745 | 750 |
| Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala | | |
| 755 | 760 | 765 |
| Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu | | |
| 770 | 775 | 780 |
| Leu Gln Val Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala | | |
| 785 | 790 | 795 |
| Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro | | |
| 805 | 810 | 815 |
| Leu Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu | | |
| 820 | 825 | 830 |
| Ser Ala Lys Gly His His His His His His | | |
| 835 | 840 | |
| <210> 51 | | |
| <211> 842 | | |
| <212> PRT | | |
| <213> Artificial | | |
| <220> | | |
| <223> Synthetic | | |
| <400> 51 | | |
| Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val | | |
| 1 | 5 | 10 |
| 15 | | |
| Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu | | |
| 20 | 25 | 30 |
| Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly | | |
| 35 | 40 | 45 |
| Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala | | |
| 50 | 55 | 60 |
| Val Phe Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala | | |
| 65 | 70 | 75 |
| 80 | | |
| Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro | | |
| 85 | 90 | 95 |
| Arg Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr | | |
| 100 | 105 | 110 |
| Arg Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu | | |
| 115 | 120 | 125 |
| Ala Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala | | |
| 130 | 135 | 140 |

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asp | Arg | Asp | Leu | Tyr | Gln | Leu | Val | Ser | Asp | Arg | Val | Ala | Val | Leu | His | 145 | 150 | 155 | 160 |
| Pro | Glu | Gly | His | Leu | Ile | Thr | Pro | Glu | Trp | Leu | Trp | Glu | Lys | Tyr | Gly | 165 | 170 | 175 | |
| Leu | Arg | Pro | Glu | Gln | Trp | Val | Asp | Phe | Arg | Ala | Leu | Val | Gly | Asp | Pro | 180 | 185 | 190 | |
| Ser | Asp | Asn | Leu | Pro | Gly | Val | Lys | Gly | Ile | Gly | Glu | Tyr | Thr | Ala | Leu | 195 | 200 | 205 | |
| Lys | Leu | Leu | Lys | Glu | Trp | Gly | Ser | Leu | Glu | Asn | Leu | Leu | Lys | Asn | Leu | 210 | 215 | 220 | |
| Asp | Arg | Val | Lys | Pro | Glu | Asn | Val | Arg | Glu | Lys | Ile | Lys | Ala | His | Leu | 225 | 230 | 235 | 240 |
| Glu | Asp | Leu | Arg | Leu | Ser | Leu | Glu | Leu | Ser | Arg | Val | Arg | Thr | Asp | Leu | 245 | 250 | 255 | |
| Pro | Leu | Glu | Val | Asp | Leu | Ala | Gln | Gly | Arg | Glu | Pro | Asp | Arg | Glu | Gly | 260 | 265 | 270 | |
| Leu | Arg | Ala | Phe | Leu | Glu | Arg | Leu | Glu | Phe | Gly | Ser | Leu | Leu | His | Glu | 275 | 280 | 285 | |
| Phe | Gly | Leu | Leu | Glu | Ala | Pro | Ala | Pro | Leu | Glu | Glu | Ala | Pro | Trp | Pro | 290 | 295 | 300 | |
| Pro | Pro | Glu | Gly | Ala | Phe | Val | Gly | Phe | Val | Leu | Ser | Arg | Pro | Glu | Pro | 305 | 310 | 315 | 320 |
| Met | Trp | Ala | Glu | Leu | Lys | Ala | Leu | Ala | Ala | Cys | Arg | Gly | Gly | Arg | Val | 325 | 330 | 335 | |
| His | Arg | Ala | Ala | Asp | Pro | Leu | Ala | Gly | Leu | Lys | Asp | Leu | Lys | Glu | Val | 340 | 345 | 350 | |
| Arg | Gly | Leu | Leu | Ala | Lys | Asp | Leu | Ala | Val | Leu | Ala | Ser | Arg | Glu | Gly | 355 | 360 | 365 | |
| Leu | Asp | Leu | Val | Pro | Gly | Asp | Asp | Pro | Met | Leu | Leu | Ala | Tyr | Leu | Leu | 370 | 375 | 380 | |
| Asp | Pro | Ser | Asn | Thr | Thr | Pro | Glu | Gly | Val | Ala | Arg | Arg | Tyr | Gly | Gly | 385 | 390 | 395 | 400 |
| Glu | Trp | Thr | Glu | Asp | Ala | Ala | His | Arg | Ala | Leu | Leu | Ser | Glu | Arg | Leu | 405 | 410 | 415 | |
| His | Arg | Asn | Leu | Leu | Lys | Arg | Leu | Glu | Gly | Glu | Glu | Lys | Leu | Leu | Trp | 420 | 425 | 430 | |
| Leu | Tyr | His | Glu | Val | Glu | Lys | Pro | Leu | Ser | Arg | Val | Leu | Ala | His | Met | 435 | 440 | 445 | |

Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala
755 760 765

Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu
770 775 780

Leu Gln Val Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala
785 790 795 800

Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro
805 810 815

Leu Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu
820 825 830

Ser Ala Lys Gly His His His His His His
835 840

<210> 52
<211> 842
<212> PRT
<213> Artificial

<220>
<223> Synthetic

<400> 52

Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
1 5 10 15

Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu
20 25 30

Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
35 40 45

Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala
50 55 60

Val Phe Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala
65 70 75 80

Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro
85 90 95

Arg Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr
100 105 110

Arg Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu
115 120 125

Ala Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala
130 135 140

Asp Arg Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His
145 150 155 160

Leu Glu Leu Ala Glu Glu Ile Arg Arg Leu Glu Glu Glu Val Phe Arg
 465 470 475 480
 Leu Ala Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg
 485 490 495
 Val Leu Phe Asp Glu Leu Arg Leu Pro Ala Leu Gly Lys Thr Gln Lys
 500 505 510
 Thr Gly Lys Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu
 515 520 525
 Ala His Pro Ile Val Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys
 530 535 540
 Leu Lys Asn Thr Tyr Val Asp Pro Leu Pro Ser Leu Val His Pro Arg
 545 550 555 560
 Thr Gly Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly
 565 570 575
 Arg Leu Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr
 580 585 590
 Pro Leu Gly Gln Arg Ile Arg Arg Ala Phe Val Ala Glu Ala Gly Trp
 595 600 605
 Ala Leu Val Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala
 610 615 620
 His Leu Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys
 625 630 635 640
 Asp Ile His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu
 645 650 655
 Ala Val Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly
 660 665 670
 Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile
 675 680 685
 Pro Tyr Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe
 690 695 700
 Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys
 705 710 715 720
 Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp
 725 730 735
 Leu Asn Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala
 740 745 750
 Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala
 755 760 765

Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu
770 775 780

Leu Gln Val Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala
785 790 795 800

Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro
805 810 815

Leu Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu
820 825 830

Ser Ala Lys Gly His His His His His His
835 840

<210> 53
<211> 839
<212> PRT
<213> Artificial

<220>
<223> Synthetic

<400> 53

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1 5 10 15

Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
20 25 30

Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
35 40 45

Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
50 55 60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
65 70 75 80

Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
85 90 95

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr Arg Leu
100 105 110

Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
115 120 125

Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
130 135 140

Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
145 150 155 160

Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|--|--|-----|--|--|--|--|--|--|--|--|--|--|--|-----|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|
| 165 | | | | | | | | | | | | | | | | 170 | | | | | | | | | | | | | | | | 175 | | | | | | | | | | | | | | | |
| Pro | Asp | Gln | Trp | Ala | Asp | Tyr | Arg | Ala | Leu | Thr | Gly | Asp | Glu | Ser | Asp | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 180 | | | | | | | | 185 | | | | | | | | 190 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Asn | Leu | Pro | Gly | Val | Lys | Gly | Ile | Gly | Glu | Lys | Thr | Ala | Arg | Lys | Leu | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 195 | | | | | | | | 200 | | | | | | | | 205 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Leu | Glu | Glu | Trp | Gly | Ser | Leu | Glu | Ala | Leu | Leu | Lys | Asn | Leu | Asp | Arg | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 210 | | | | | | | | 215 | | | | | | | | 220 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Leu | Lys | Pro | Ala | Ile | Arg | Glu | Lys | Ile | Leu | Ala | His | Met | Asp | Asp | Leu | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 225 | | | | | | | | 230 | | | | | | | | 235 | | | | 240 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Lys | Leu | Ser | Trp | Asp | Leu | Ala | Lys | Val | Arg | Thr | Asp | Leu | Pro | Leu | Glu | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | 245 | | | | | | | | 250 | | | | | | | | 255 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Val | Asp | Phe | Ala | Lys | Arg | Arg | Glu | Pro | Asp | Arg | Glu | Arg | Leu | Arg | Ala | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | 260 | | | | | | | | 265 | | | | | | | | 270 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Phe | Leu | Glu | Arg | Leu | Glu | Phe | Gly | Ser | Leu | Leu | His | Glu | Phe | Gly | Leu | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | 275 | | | | | | | | 280 | | | | | | | | 285 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Leu | Glu | Ser | Pro | Lys | Ala | Leu | Glu | Glu | Ala | Pro | Trp | Pro | Pro | Pro | Glu | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | 290 | | | | | | | | 295 | | | | | | | | 300 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Gly | Ala | Phe | Val | Gly | Phe | Val | Leu | Ser | Arg | Lys | Glu | Pro | Met | Trp | Ala | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 305 | | | | | | | | 310 | | | | | | | | 315 | | | | 320 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Asp | Leu | Leu | Ala | Leu | Ala | Ala | Ala | Arg | Gly | Gly | Arg | Val | His | Arg | Ala | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | 325 | | | | | | | | 330 | | | | | | | | 335 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Pro | Glu | Pro | Tyr | Lys | Ala | Leu | Arg | Asp | Leu | Lys | Glu | Ala | Arg | Gly | Leu | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | 340 | | | | | | | | 345 | | | | | | | | 350 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Leu | Ala | Lys | Asp | Leu | Ser | Val | Leu | Ala | Leu | Arg | Glu | Gly | Leu | Gly | Leu | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | 355 | | | | | | | | 360 | | | | | | | | 365 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Pro | Pro | Gly | Asp | Asp | Pro | Met | Leu | Leu | Ala | Tyr | Leu | Leu | Asp | Pro | Ser | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 370 | | | | | | | | 375 | | | | | | | | 380 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Asn | Thr | Thr | Pro | Glu | Gly | Val | Ala | Arg | Arg | Tyr | Gly | Gly | Glu | Trp | Thr | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 385 | | | | | | | | 390 | | | | | | | | 395 | | | | 400 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Glu | Glu | Ala | Gly | Glu | Arg | Ala | Ala | Leu | Ser | Glu | Arg | Leu | Phe | Ala | Asn | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | 405 | | | | | | | | 410 | | | | | | | | 415 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Leu | Leu | Lys | Arg | Leu | Glu | Gly | Glu | Glu | Arg | Leu | Leu | Trp | Leu | Tyr | Arg | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | 420 | | | | | | | | 425 | | | | | | | | 430 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Glu | Val | Glu | Arg | Pro | Leu | Ser | Ala | Val | Leu | Ala | His | Met | Glu | Ala | Thr | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | 435 | | | | | | | | 440 | | | | | | | | 445 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Gly | Val | Arg | Leu | Asp | Val | Ala | Tyr | Leu | Arg | Ala | Leu | Ser | Leu | Glu | Val | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 450 | | | | | | | | 455 | | | | | | | | 460 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Ala | Glu | Glu | Ile | Ala | Arg | Leu | Glu | Ala | Glu | Val | Phe | Arg | Leu | Ala | Gly | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

| | | |
|---------------------|---------------------|-------------------------|
| 770 | 775 | 780 |
| Ala Asn Glu Leu Val | Leu Glu Ala Pro Lys | Glu Arg Ala Glu Ala Val |
| 785 | 790 | 795 800 |
| Ala Arg Leu Ala Lys | Glu Val Met Glu Gly | Val Tyr Pro Leu Ala Val |
| | 805 | 810 815 |
| Pro Leu Glu Val Glu | Val Gly Ile Gly Glu | Asp Trp Leu Ser Ala Lys |
| | 820 | 825 830 |
| Glu His His His His | His His | |
| | 835 | |
| <210> 54 | | |
| <211> 839 | | |
| <212> PRT | | |
| <213> Artificial | | |
| <220> | | |
| <223> Synthetic | | |
| <400> 54 | | |
| Met Asn Ser Gly Met | Leu Pro Leu Phe Glu | Pro Lys Gly Arg Val Leu |
| 1 | 5 | 10 15 |
| Leu Val Asp Gly His | His Leu Ala Tyr Arg | Thr Phe His Ala Leu Lys |
| | 20 | 25 30 |
| Gly Leu Thr Thr Ser | Arg Gly Glu Pro Val | Gln Ala Val Tyr Gly Phe |
| | 35 | 40 45 |
| Ala Lys Ser Leu Leu | Lys Ala Leu Lys Glu | Asp Gly Asp Ala Val Ile |
| | 50 | 55 60 |
| Val Val Phe Asp Ala | Lys Ala Pro Ser Phe | Arg His Glu Ala Tyr Gly |
| 65 | 70 | 75 80 |
| Gly Tyr Lys Ala Gly | Arg Ala Glu Thr Glu | Glu Asp Phe Pro Arg Gln |
| | 85 | 90 95 |
| Leu Ala Leu Ile Lys | Glu Leu Val Asp Leu | Leu Gly Leu Ala Arg Leu |
| | 100 | 105 110 |
| Glu Val Pro Gly Tyr | Glu Ala Asp Asp Val | Leu Ala Ser Leu Ala Lys |
| | 115 | 120 125 |
| Lys Ala Glu Lys Glu | Gly Tyr Glu Val Arg | Ile Leu Thr Ala Asp Lys |
| | 130 | 135 140 |
| Asp Leu Tyr Gln Leu | Leu Ser Asp Arg Ile | His Val Leu His Pro Glu |
| 145 | 150 | 155 160 |
| Gly Tyr Leu Ile Thr | Pro Ala Trp Leu Trp | Glu Lys Tyr Gly Leu Arg |
| | 165 | 170 175 |

Asn Leu Pro Gly Val Asn Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
 195 200 205
 Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
 210 215 220
 Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
 225 230 235 240
 Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
 245 250 255
 Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
 260 265 270
 Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
 275 280 285
 Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu
 290 295 300
 Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala
 305 310 315 320
 Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala
 325 330 335
 Pro Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu
 340 345 350
 Leu Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu
 355 360 365
 Pro Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser
 370 375 380
 Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr
 385 390 395 400
 Glu Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn
 405 410 415
 Leu Leu Lys Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg
 420 425 430
 Glu Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr
 435 440 445
 Gly Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val
 450 455 460
 Ala Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly
 465 470 475 480
 His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe
 485 490 495

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Glu His His His His His His
835

<210> 56

<211> 839

<212> PRT

<213> Artificial

<220>

<223> Synthetic

<400> 56

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1 5 10 15

Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
20 25 30

Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
35 40 45

Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
50 55 60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
65 70 75 80

Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
85 90 95

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr Arg Leu
100 105 110

Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
115 120 125

Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
130 135 140

Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
145 150 155 160

Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
165 170 175

Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
180 185 190

Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Gln Arg Lys Leu

| | | |
|-------------------------|-------------------------|---------------------|
| 195 | 200 | 205 |
| Leu Glu Glu Trp Gly Ser | Leu Glu Ala Leu Leu Lys | Asn Leu Asp Arg |
| 210 | 215 | 220 |
| Leu Lys Pro Ala Ile Arg | Glu Lys Ile Leu Ala His | Met Asp Asp Leu |
| 225 | 230 | 235 |
| Lys Leu Ser Trp Asp | Leu Ala Lys Val Arg Thr | Asp Leu Pro Leu Glu |
| 245 | 250 | 255 |
| Val Asp Phe Ala Lys Arg | Arg Glu Pro Asp Arg | Glu Arg Leu Arg Ala |
| 260 | 265 | 270 |
| Phe Leu Glu Arg Leu Glu | Phe Gly Ser Leu Leu His | Glu Phe Gly Leu |
| 275 | 280 | 285 |
| Leu Glu Ser Pro Lys Ala | Leu Glu Glu Ala Pro Trp | Pro Pro Pro Glu |
| 290 | 295 | 300 |
| Gly Ala Phe Val Gly Phe | Val Leu Ser Arg Lys | Glu Pro Met Trp Ala |
| 305 | 310 | 315 |
| Asp Leu Leu Ala Leu Ala | Ala Ala Arg Gly Gly | Arg Val His Arg Ala |
| 325 | 330 | 335 |
| Pro Glu Pro Tyr Lys Ala | Leu Arg Asp Leu Lys | Glu Ala Arg Gly Leu |
| 340 | 345 | 350 |
| Leu Ala Lys Asp Leu Ser | Val Leu Ala Leu Arg | Glu Gly Leu Gly Leu |
| 355 | 360 | 365 |
| Pro Pro Gly Asp Asp Pro | Met Leu Leu Ala Tyr | Leu Leu Asp Pro Ser |
| 370 | 375 | 380 |
| Asn Thr Thr Pro Glu Gly | Val Ala Arg Arg Tyr | Gly Gly Glu Trp Thr |
| 385 | 390 | 395 |
| Glu Glu Ala Gly Glu Arg | Ala Ala Leu Ser Glu | Arg Leu Phe Ala Asn |
| 405 | 410 | 415 |
| Leu Leu Lys Arg Leu Glu | Gly Glu Glu Arg Leu | Leu Trp Leu Tyr Arg |
| 420 | 425 | 430 |
| Glu Val Glu Arg Pro Leu | Ser Ala Val Leu Ala | His Met Glu Ala Thr |
| 435 | 440 | 445 |
| Gly Val Arg Leu Asp Val | Ala Tyr Leu Arg Ala | Leu Ser Leu Glu Val |
| 450 | 455 | 460 |
| Ala Glu Glu Ile Ala Arg | Leu Glu Ala Glu Val | Phe Arg Leu Ala Gly |
| 465 | 470 | 475 |
| His Pro Phe Asn Leu Asn | Ser Arg Asp Gln Leu | Glu Arg Val Leu Phe |
| 485 | 490 | 495 |
| Asp Glu Leu Arg Leu Pro | Lys Leu Lys Lys Thr | Lys Lys Thr Gly Lys |

| | | |
|---|-----|-----|
| 500 | 505 | 510 |
| Arg Ser Ser Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro | | |
| 515 | 520 | 525 |
| Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser | | |
| 530 | 535 | 540 |
| Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg | | |
| 545 | 550 | 555 |
| Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser | | |
| 565 | 570 | 575 |
| Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly | | |
| 580 | 585 | 590 |
| Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val | | |
| 595 | 600 | 605 |
| Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser | | |
| 610 | 615 | 620 |
| Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His | | |
| 625 | 630 | 635 |
| Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp | | |
| 645 | 650 | 655 |
| Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr | | |
| 660 | 665 | 670 |
| Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu | | |
| 675 | 680 | 685 |
| Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val | | |
| 690 | 695 | 700 |
| Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr | | |
| 705 | 710 | 715 |
| Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala | | |
| 725 | 730 | 735 |
| Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met | | |
| 740 | 745 | 750 |
| Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys | | |
| 755 | 760 | 765 |
| Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val | | |
| 770 | 775 | 780 |
| Ala Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val | | |
| 785 | 790 | 795 |
| Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val | | |

| | | | | | | | | | | | | | | | | |
|------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Leu 225 | Lys | Pro | Ala | Ile | Arg | Glu | Lys | Ile | Leu | Ala | His | Met | Asp | Asp | Leu | 240 |
| Lys | Leu | Ser | Trp | Asp | Leu | Ala | Lys | Val | Arg | Thr | Asp | Leu | Pro | Leu | Glu | 245 |
| Val | Asp | Phe | Ala | Lys | Arg | Arg | Glu | Pro | Asp | Arg | Glu | Arg | Leu | Arg | Ala | 250 |
| Phe | Leu | Glu | Arg | Leu | Glu | Phe | Gly | Ser | Leu | Leu | His | Glu | Phe | Gly | Leu | 255 |
| Leu | Glu | Ser | Pro | Lys | Ala | Leu | Glu | Glu | Ala | Pro | Trp | Pro | Pro | Pro | Glu | 260 |
| Gly | Ala | Phe | Val | Gly | Phe | Val | Leu | Ser | Arg | Lys | Glu | Pro | Met | Trp | Ala | 265 |
| Asp | Leu | Leu | Ala | Leu | Ala | Ala | Ala | Arg | Gly | Gly | Arg | Val | His | Arg | Ala | 270 |
| Pro | Glu | Pro | Tyr | Lys | Ala | Leu | Arg | Asp | Leu | Lys | Glu | Ala | Arg | Gly | Leu | 275 |
| Leu | Ala | Lys | Asp | Leu | Ser | Val | Leu | Ala | Leu | Arg | Glu | Gly | Leu | Gly | Leu | 280 |
| Pro | Pro | Gly | Asp | Asp | Pro | Met | Leu | Leu | Ala | Tyr | Leu | Leu | Asp | Pro | Ser | 285 |
| Asn | Thr | Thr | Pro | Glu | Gly | Val | Ala | Arg | Arg | Tyr | Gly | Gly | Glu | Trp | Thr | 290 |
| Glu | Glu | Ala | Gly | Glu | Arg | Ala | Ala | Leu | Ser | Glu | Arg | Leu | Phe | Ala | Asn | 295 |
| Leu | Leu | Lys | Arg | Leu | Glu | Gly | Glu | Glu | Arg | Leu | Leu | Trp | Leu | Tyr | Arg | 300 |
| Glu | Val | Glu | Arg | Pro | Leu | Ser | Ala | Val | Leu | Ala | His | Met | Glu | Ala | Thr | 305 |
| Gly | Val | Arg | Leu | Asp | Val | Ala | Tyr | Leu | Arg | Ala | Leu | Ser | Leu | Glu | Val | 310 |
| Ala | Glu | Glu | Ile | Ala | Arg | Leu | Glu | Ala | Glu | Val | Phe | Arg | Leu | Ala | Gly | 315 |
| His | Pro | Phe | Asn | Leu | Asn | Ser | Arg | Asp | Gln | Leu | Glu | Arg | Val | Leu | Phe | 320 |
| Asp | Glu | Leu | Arg | Leu | Pro | Lys | Leu | Lys | Lys | Thr | Lys | Lys | Thr | Gly | Lys | 325 |
| Arg | Ser | Ser | Ser | Ala | Ala | Val | Leu | Glu | Ala | Leu | Arg | Glu | Ala | His | Pro | 330 |

| | | | |
|---|-----|-----|---------|
| Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser | 530 | 535 | 540 |
| Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg | 545 | 550 | 555 560 |
| Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser | 565 | 570 | 575 |
| Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly | 580 | 585 | 590 |
| Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val | 595 | 600 | 605 |
| Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser | 610 | 615 | 620 |
| Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His | 625 | 630 | 635 640 |
| Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp | 645 | 650 | 655 |
| Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr | 660 | 665 | 670 |
| Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu | 675 | 680 | 685 |
| Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val | 690 | 695 | 700 |
| Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr | 705 | 710 | 715 720 |
| Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala | 725 | 730 | 735 |
| Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met | 740 | 745 | 750 |
| Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys | 755 | 760 | 765 |
| Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val | 770 | 775 | 780 |
| Ala Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val | 785 | 790 | 795 800 |
| Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val | 805 | 810 | 815 |
| Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys | 820 | 825 | 830 |

Glu His His His His His His
835

<210> 59
<211> 839
<212> PRT
<213> Artificial

<220>
<223> Synthetic

<400> 59

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1 5 10 15
Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
20 25 30
Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
35 40 45
Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
50 55 60
Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
65 70 75 80
Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
85 90 95
Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr Arg Leu
100 105 110
Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
115 120 125
Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
130 135 140
Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
145 150 155 160
Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
165 170 175
Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
180 185 190
Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Pro Ala Arg Lys Leu
195 200 205
Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
210 215 220
Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu

| | | | | | | |
|---|---|-----|-----|-----|-----|-----|
| 225 | | 230 | | 235 | | 240 |
| Lys Leu Ser Trp Asp | Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu | | | | | |
| | 245 | | 250 | | 255 | |
| Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala | | | | | | |
| | 260 | | 265 | | 270 | |
| Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu | | | | | | |
| | 275 | | 280 | | 285 | |
| Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu | | | | | | |
| | 290 | | 295 | | 300 | |
| Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala | | | | | | |
| 305 | | 310 | | 315 | | 320 |
| Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala | | | | | | |
| | 325 | | 330 | | 335 | |
| Pro Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu | | | | | | |
| | 340 | | 345 | | 350 | |
| Leu Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu | | | | | | |
| | 355 | | 360 | | 365 | |
| Pro Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser | | | | | | |
| | 370 | | 375 | | 380 | |
| Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr | | | | | | |
| 385 | | 390 | | 395 | | 400 |
| Glu Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn | | | | | | |
| | 405 | | 410 | | 415 | |
| Leu Leu Lys Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg | | | | | | |
| | 420 | | 425 | | 430 | |
| Glu Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr | | | | | | |
| | 435 | | 440 | | 445 | |
| Gly Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val | | | | | | |
| | 450 | | 455 | | 460 | |
| Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly | | | | | | |
| 465 | | 470 | | 475 | | 480 |
| His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe | | | | | | |
| | 485 | | 490 | | 495 | |
| Asp Glu Leu Arg Leu Pro Lys Leu Lys Lys Thr Lys Lys Thr Gly Lys | | | | | | |
| | 500 | | 505 | | 510 | |
| Arg Ser Ser Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro | | | | | | |
| | 515 | | 520 | | 525 | |
| Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser | | | | | | |
| | 530 | | 535 | | 540 | |

<210> 60
 <211> 839
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 60

```

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1          5          10          15

Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
          20          25          30

Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
          35          40          45

Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
          50          55          60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
65          70          75          80

Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
          85          90          95

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr Arg Leu
          100          105          110

Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
          115          120          125

Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
          130          135          140

Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
145          150          155          160

Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
          165          170          175

Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
          180          185          190

Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
          195          200          205

Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
          210          215          220

Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
225          230          235          240

Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu

```

| | | | | | |
|---|-----|--|-----|--|-----|
| | 245 | | 250 | | 255 |
| Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala | 260 | | 265 | | 270 |
| Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu | 275 | | 280 | | 285 |
| Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu | 290 | | 295 | | 300 |
| Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala | 305 | | 310 | | 315 |
| Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala | 325 | | 330 | | 335 |
| Pro Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu | 340 | | 345 | | 350 |
| Leu Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu | 355 | | 360 | | 365 |
| Pro Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser | 370 | | 375 | | 380 |
| Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr | 385 | | 390 | | 395 |
| Glu Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn | 405 | | 410 | | 415 |
| Leu Leu Lys Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg | 420 | | 425 | | 430 |
| Glu Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr | 435 | | 440 | | 445 |
| Gly Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val | 450 | | 455 | | 460 |
| Ala Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly | 465 | | 470 | | 475 |
| His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe | 485 | | 490 | | 495 |
| Asp Glu Leu Arg Leu Pro Lys Leu Lys Lys Thr Lys Lys Thr Gly Lys | 500 | | 505 | | 510 |
| Arg Ser Ser Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro | 515 | | 520 | | 525 |
| Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser | 530 | | 535 | | 540 |
| Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg | | | | | |

| Variable | Mean | SD | Min | Max |
|----------------------|------|------|-----|-----|
| Age | 34.5 | 10.2 | 20 | 55 |
| Gender | 0.5 | 0.5 | 0 | 1 |
| Marital status | 0.6 | 0.5 | 0 | 1 |
| Education | 12.5 | 1.5 | 9 | 16 |
| Income | 15.2 | 5.8 | 5 | 30 |
| Occupation | 1.2 | 0.8 | 0 | 2 |
| Health status | 1.5 | 0.5 | 1 | 2 |
| Life satisfaction | 4.2 | 1.2 | 1 | 7 |
| Depression | 1.8 | 0.8 | 1 | 3 |
| Stress | 2.5 | 0.9 | 1 | 4 |
| Resilience | 3.5 | 1.0 | 1 | 5 |
| Optimism | 4.5 | 1.1 | 1 | 7 |
| Gratitude | 5.2 | 1.3 | 1 | 8 |
| Forgiveness | 4.8 | 1.2 | 1 | 7 |
| Empathy | 5.5 | 1.4 | 1 | 8 |
| Compassion | 5.8 | 1.5 | 1 | 9 |
| Kindness | 6.2 | 1.6 | 1 | 9 |
| Generosity | 6.5 | 1.7 | 1 | 10 |
| Patience | 6.8 | 1.8 | 1 | 10 |
| Humility | 7.2 | 1.9 | 1 | 10 |
| Modesty | 7.5 | 2.0 | 1 | 10 |
| Shyness | 7.8 | 2.1 | 1 | 10 |
| Introversion | 8.2 | 2.2 | 1 | 10 |
| Extroversion | 8.5 | 2.3 | 1 | 10 |
| Sensitivity | 8.8 | 2.4 | 1 | 10 |
| Emotionality | 9.2 | 2.5 | 1 | 10 |
| Imagination | 9.5 | 2.6 | 1 | 10 |
| Creativity | 9.8 | 2.7 | 1 | 10 |
| Innovation | 10.2 | 2.8 | 1 | 10 |
| Leadership | 10.5 | 2.9 | 1 | 10 |
| Management | 10.8 | 3.0 | 1 | 10 |
| Organization | 11.2 | 3.1 | 1 | 10 |
| Planning | 11.5 | 3.2 | 1 | 10 |
| Decision making | 11.8 | 3.3 | 1 | 10 |
| Problem solving | 12.2 | 3.4 | 1 | 10 |
| Communication | 12.5 | 3.5 | 1 | 10 |
| Interpersonal skills | 12.8 | 3.6 | 1 | 10 |
| Teamwork | 13.2 | 3.7 | 1 | 10 |
| Collaboration | 13.5 | 3.8 | 1 | 10 |
| Cooperation | 13.8 | 3.9 | 1 | 10 |
| Helpfulness | 14.2 | 4.0 | 1 | 10 |
| Supportiveness | 14.5 | 4.1 | 1 | 10 |
| Encouragement | 14.8 | 4.2 | 1 | 10 |
| Inspiration | 15.2 | 4.3 | 1 | 10 |
| Motivation | 15.5 | 4.4 | 1 | 10 |
| Goal setting | 15.8 | 4.5 | 1 | 10 |
| Time management | 16.2 | 4.6 | 1 | 10 |
| Productivity | 16.5 | 4.7 | 1 | 10 |
| Efficiency | 16.8 | 4.8 | 1 | 10 |
| Effectiveness | 17.2 | 4.9 | 1 | 10 |
| Success | 17.5 | 5.0 | 1 | 10 |
| achievement | 17.8 | 5.1 | 1 | 10 |
| Recognition | 18.2 | 5.2 | 1 | 10 |
| Respect | 18.5 | 5.3 | 1 | 10 |
| Honor | 18.8 | 5.4 | 1 | 10 |
| Glory | 19.2 | 5.5 | 1 | 10 |
| Fame | 19.5 | 5.6 | 1 | 10 |
| Power | 19.8 | 5.7 | 1 | 10 |
| Influence | 20.2 | 5.8 | 1 | 10 |
| Control | 20.5 | 5.9 | 1 | 10 |
| Authority | 20.8 | 6.0 | 1 | 10 |
| Leadership | 21.2 | 6.1 | 1 | 10 |
| Management | 21.5 | 6.2 | 1 | 10 |
| Organization | 21.8 | 6.3 | 1 | 10 |
| Planning | 22.2 | 6.4 | 1 | 10 |
| Decision making | 22.5 | 6.5 | 1 | 10 |
| Problem solving | 22.8 | 6.6 | 1 | 10 |
| Communication | 23.2 | 6.7 | 1 | 10 |
| Interpersonal skills | 23.5 | 6.8 | 1 | 10 |
| Teamwork | 23.8 | 6.9 | 1 | 10 |
| Collaboration | 24.2 | 7.0 | 1 | 10 |
| Cooperation | 24.5 | 7.1 | 1 | 10 |
| Helpfulness | 24.8 | 7.2 | 1 | 10 |
| Supportiveness | 25.2 | 7.3 | 1 | 10 |
| Encouragement | 25.5 | 7.4 | 1 | 10 |
| Inspiration | 25.8 | 7.5 | 1 | 10 |
| Motivation | 26.2 | 7.6 | 1 | 10 |
| Goal setting | 26.5 | 7.7 | 1 | 10 |
| Time management | 26.8 | 7.8 | 1 | 10 |
| Productivity | 27.2 | 7.9 | 1 | 10 |
| Efficiency | 27.5 | 8.0 | 1 | 10 |
| Effectiveness | 27.8 | 8.1 | 1 | 10 |
| Success | 28.2 | 8.2 | 1 | 10 |
| achievement | 28.5 | | | |

<400> 61

464

Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
260 265 270

Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
275 280 285

Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu
290 295 300

Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala
305 310 315 320

Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala
325 330 335

Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val Arg Gly Leu
340 345 350

Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly Leu Asp Leu
355 360 365

Val Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser
370 375 380

Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr
385 390 395 400

Glu Asp Ala Ala His Arg Ala Leu Leu Ser Glu Arg Leu His Arg Asn
405 410 415

Leu Leu Lys Arg Leu Glu Gly Glu Glu Lys Leu Leu Trp Leu Tyr His
420 425 430

Glu Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met Glu Ala Thr
435 440 445

Gly Val Arg Arg Asp Val Ala Tyr Leu Gln Ala Leu Ser Leu Glu Leu
450 455 460

Ala Glu Glu Ile Arg Arg Leu Glu Glu Glu Val Phe Arg Leu Ala Gly
465 470 475 480

His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe
485 490 495

Asp Glu Leu Arg Leu Pro Ala Leu Lys Lys Thr Lys Lys Thr Gly Lys
500 505 510

Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro
515 520 525

Ile Val Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys Leu Lys Asn
530 535 540

Thr Tyr Val Asp Pro Leu Pro Ser Leu Val His Pro Arg Thr Gly Arg
545 550 555 560

<213> Artificial

<220>

<223> Synthetic

<400> 62

Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
1 5 10 15
Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu
20 25 30
Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
35 40 45
Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala
50 55 60
Val Phe Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala
65 70 75 80
Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro
85 90 95
Arg Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr
100 105 110
Arg Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu
115 120 125
Ala Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala
130 135 140
Asp Arg Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His
145 150 155 160
Pro Glu Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly
165 170 175
Leu Arg Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro
180 185 190
Ser Asp Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu
195 200 205
Lys Leu Leu Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu
210 215 220
Asp Arg Val Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu
225 230 235 240
Glu Asp Leu Arg Leu Ser Leu Glu Leu Ser Arg Val Arg Thr Asp Leu
245 250 255
Pro Leu Glu Val Asp Leu Ala Gln Gly Arg Glu Pro Asp Arg Glu Gly
260 265 270

09064536-03404

Leu Arg Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu
 275 280 285
 Phe Gly Leu Leu Glu Ala Pro Ala Pro Leu Glu Glu Ala Pro Trp Pro
 290 295 300
 Pro Pro Glu Gly Ala Phe Val Gly Phe Val Leu Ser Arg Pro Glu Pro
 305 310 315 320
 Met Trp Ala Glu Leu Lys Ala Leu Ala Ala Cys Arg Gly Gly Arg Val
 325 330 335
 His Arg Ala Pro Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala
 340 345 350
 Arg Gly Leu Leu Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly
 355 360 365
 Leu Gly Leu Pro Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu
 370 375 380
 Asp Pro Ser Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly
 385 390 395 400
 Glu Trp Thr Glu Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu
 405 410 415
 Phe Ala Asn Leu Leu Lys Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp
 420 425 430
 Leu Tyr Arg Glu Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met
 435 440 445
 Glu Ala Thr Gly Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser
 450 455 460
 Leu Glu Val Ala Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg
 465 470 475 480
 Leu Ala Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg
 485 490 495
 Val Leu Phe Asp Glu Leu Gly Leu Pro Ala Ile Lys Lys Thr Gln Lys
 500 505 510
 Thr Gly Lys Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu
 515 520 525
 Ala His Pro Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys
 530 535 540
 Leu Lys Ser Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg
 545 550 555 560
 Thr Gly Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly
 565 570 575

| Variable | Mean | SD | Min | Max |
|-------------------------|------|------|-----|-----|
| Age | 35.2 | 12.5 | 18 | 65 |
| Gender | 50.0 | 0.0 | 0 | 1 |
| Marital status | 65.0 | 15.0 | 0 | 100 |
| Education | 12.5 | 2.0 | 8 | 16 |
| Income | 30.0 | 10.0 | 10 | 50 |
| Health status | 75.0 | 10.0 | 50 | 100 |
| Life satisfaction | 60.0 | 15.0 | 30 | 90 |
| Stress level | 40.0 | 12.0 | 20 | 60 |
| Work-life balance | 55.0 | 18.0 | 30 | 80 |
| Family support | 70.0 | 12.0 | 50 | 90 |
| Community involvement | 60.0 | 15.0 | 40 | 80 |
| Personal growth | 50.0 | 10.0 | 30 | 70 |
| Financial stability | 65.0 | 12.0 | 45 | 85 |
| Emotional well-being | 70.0 | 10.0 | 50 | 90 |
| Physical health | 80.0 | 8.0 | 60 | 100 |
| Mental health | 75.0 | 10.0 | 55 | 95 |
| Social relationships | 65.0 | 12.0 | 45 | 85 |
| Work satisfaction | 55.0 | 15.0 | 35 | 75 |
| Life goals achievement | 60.0 | 18.0 | 40 | 80 |
| Personal happiness | 65.0 | 10.0 | 45 | 85 |
| Overall quality of life | 60.0 | 15.0 | 40 | 80 |

470

Phe Gly Leu Leu Gly Gly Glu Lys Pro Arg Glu Glu Ala Pro Trp Pro
 290 295 300
 Pro Pro Glu Gly Ala Phe Val Gly Phe Leu Leu Ser Arg Lys Glu Pro
 305 310 315 320
 Met Trp Ala Glu Leu Leu Ala Leu Ala Ala Ala Ser Glu Gly Arg Val
 325 330 335
 His Arg Ala Thr Ser Pro Val Glu Ala Leu Ala Asp Leu Lys Glu Ala
 340 345 350
 Arg Gly Phe Leu Ala Lys Asp Leu Ala Val Leu Ala Leu Arg Glu Gly
 355 360 365
 Val Ala Leu Asp Pro Thr Asp Asp Pro Leu Leu Val Ala Tyr Leu Leu
 370 375 380
 Asp Pro Ala Asn Thr His Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly
 385 390 395 400
 Glu Phe Thr Glu Asp Ala Ala Glu Arg Ala Leu Leu Ser Glu Arg Leu
 405 410 415
 Phe Gln Asn Leu Phe Lys Arg Leu Ser Glu Lys Leu Leu Trp Leu Tyr
 420 425 430
 Gln Glu Val Glu Arg Pro Leu Ser Arg Val Leu Ala His Met Glu Ala
 435 440 445
 Arg Gly Val Arg Leu Asp Val Pro Leu Leu Glu Ala Leu Ser Phe Glu
 450 455 460
 Leu Glu Lys Glu Met Glu Arg Leu Glu Gly Glu Val Phe Arg Leu Ala
 465 470 475 480
 Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu
 485 490 495
 Phe Asp Glu Leu Gly Leu Thr Pro Val Gly Arg Thr Gln Lys Thr Gly
 500 505 510
 Lys Arg Ser Thr Ala Gln Gly Ala Leu Glu Ala Leu Arg Gly Ala His
 515 520 525
 Pro Ile Val Glu Leu Ile Leu Gln Tyr Arg Glu Leu Ser Lys Leu Lys
 530 535 540
 Ser Thr Tyr Leu Asp Pro Leu Pro Arg Leu Val His Pro Arg Thr Gly
 545 550 555 560
 Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu
 565 570 575
 Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu
 580 585 590

<400> 64

Met Glu Phe Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu Val
1 5 10 15
Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu Lys Gly Leu
20 25 30
Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala Lys
35 40 45
Ser Leu Leu Lys Ala Leu Arg Glu Asp Gly Asp Val Val Ile Val Val
50 55 60
Phe Asp Ala Lys Ala Pro Ser Phe Arg His Gln Thr Tyr Glu Ala Tyr
65 70 75 80
Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu Ala
85 90 95
Leu Ile Lys Glu Met Val Asp Leu Leu Gly Leu Glu Arg Leu Glu Val
100 105 110
Pro Gly Phe Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys Lys Ala
115 120 125
Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Arg Asp Leu
130 135 140
Tyr Gln Leu Leu Ser Glu Arg Ile Ser Ile Leu His Pro Glu Gly Tyr
145 150 155 160
Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly Leu Lys Pro Ser
165 170 175
Gln Trp Val Asp Tyr Arg Ala Leu Ala Gly Asp Pro Ser Asp Asn Ile
180 185 190
Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Ala Lys Leu Ile Arg
195 200 205
Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys His Leu Glu Gln Val Lys
210 215 220
Pro Ala Ser Val Arg Glu Lys Ile Leu Ser His Met Glu Asp Leu Lys
225 230 235 240
Leu Ser Leu Glu Leu Ser Arg Val His Thr Asp Leu Leu Leu Gln Val
245 250 255
Asp Phe Ala Arg Arg Arg Glu Pro Asp Arg Glu Gly Leu Lys Ala Phe
260 265 270
Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu Leu
275 280 285
Glu Ser Pro Val Ala Ala Glu Glu Ala Pro Trp Pro Pro Pro Glu Gly

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Leu | Leu | Val | Asp | Gly | His | His | Leu | Ala | Tyr | Arg | Thr | Phe | Tyr | Ala | Leu |
| | | 20 | | | | | | 25 | | | | | 30 | | |
| Ser | Leu | Thr | Thr | Ser | Arg | Gly | Glu | Pro | Val | Gln | Met | Val | Tyr | Gly | Phe |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Ala | Arg | Ser | Leu | Leu | Lys | Ala | Leu | Lys | Glu | Asp | Gly | Gln | Ala | Val | Val |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| Val | Val | Phe | Asp | Ala | Lys | Ala | Pro | Ser | Phe | Arg | His | Glu | Ala | Tyr | Glu |
| 65 | | | | | 70 | | | | | 75 | | | | | 80 |
| Ala | Tyr | Lys | Ala | Gly | Arg | Ala | Pro | Thr | Pro | Glu | Asp | Phe | Pro | Arg | Gln |
| | | | | 85 | | | | | 90 | | | | | 95 | |
| Leu | Ala | Leu | Val | Lys | Arg | Leu | Val | Asp | Leu | Leu | Gly | Leu | Val | Arg | Leu |
| | | | 100 | | | | | 105 | | | | | 110 | | |
| Glu | Ala | Pro | Gly | Tyr | Glu | Ala | Asp | Asp | Val | Leu | Gly | Thr | Leu | Ala | Lys |
| | | 115 | | | | | 120 | | | | | 125 | | | |
| Lys | Ala | Glu | Arg | Glu | Gly | Met | Glu | Val | Arg | Ile | Leu | Thr | Gly | Asp | Arg |
| | 130 | | | | | 135 | | | | | 140 | | | | |
| Asp | Phe | Phe | Gln | Leu | Leu | Ser | Glu | Lys | Val | Ser | Val | Leu | Leu | Pro | Asp |
| 145 | | | | | 150 | | | | | 155 | | | | | 160 |
| Gly | Thr | Leu | Val | Thr | Pro | Lys | Asp | Val | Gln | Glu | Lys | Tyr | Gly | Val | Pro |
| | | | | 165 | | | | | 170 | | | | | 175 | |
| Pro | Glu | Arg | Trp | Val | Asp | Phe | Arg | Ala | Leu | Thr | Gly | Asp | Arg | Ser | Asp |
| | | | 180 | | | | | 185 | | | | | 190 | | |
| Asn | Ile | Pro | Gly | Val | Ala | Gly | Ile | Gly | Glu | Lys | Thr | Ala | Leu | Arg | Leu |
| | | 195 | | | | | 200 | | | | | 205 | | | |
| Leu | Ala | Glu | Trp | Gly | Ser | Val | Glu | Asn | Leu | Leu | Lys | Asn | Leu | Asp | Arg |
| | 210 | | | | | 215 | | | | | 220 | | | | |
| Val | Lys | Pro | Asp | Ser | Leu | Arg | Arg | Lys | Ile | Glu | Ala | His | Leu | Glu | Asp |
| 225 | | | | | 230 | | | | | 235 | | | | | 240 |
| Leu | His | Leu | Ser | Leu | Asp | Leu | Ala | Arg | Ile | Arg | Thr | Asp | Leu | Pro | Leu |
| | | | | 245 | | | | | 250 | | | | 255 | | |
| Glu | Val | Asp | Phe | Lys | Ala | Leu | Arg | Arg | Arg | Thr | Pro | Asp | Leu | Glu | Gly |
| | | | 260 | | | | | 265 | | | | | 270 | | |
| Leu | Arg | Ala | Phe | Leu | Glu | Glu | Leu | Glu | Phe | Gly | Ser | Leu | Leu | His | Glu |
| | | 275 | | | | | 280 | | | | | 285 | | | |
| Phe | Gly | Leu | Leu | Gly | Gly | Glu | Lys | Pro | Arg | Glu | Glu | Ala | Pro | Trp | Pro |
| | 290 | | | | | 295 | | | | | 300 | | | | |
| Pro | Pro | Glu | Gly | Ala | Phe | Val | Gly | Phe | Leu | Leu | Ser | Arg | Lys | Glu | Pro |
| 305 | | | | | 310 | | | | | | 315 | | | | 320 |

| | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| Met | Trp | Ala | Glu | Leu | Leu | Ala | Leu | Ala | Ala | Ala | Ser | Gly | Gly | Arg | Val | |
| | | | | 325 | | | | | 330 | | | | | | 335 | |
| His | Arg | Ala | Ala | Asp | Pro | Leu | Ala | Gly | Leu | Lys | Asp | Leu | Lys | Glu | Val | |
| | | | 340 | | | | | 345 | | | | | 350 | | | |
| Arg | Gly | Leu | Leu | Ala | Lys | Asp | Leu | Ala | Val | Leu | Ala | Ser | Arg | Glu | Gly | |
| | | 355 | | | | | 360 | | | | | 365 | | | | |
| Leu | Asp | Leu | Val | Pro | Gly | Asp | Asp | Pro | Met | Leu | Leu | Ala | Tyr | Leu | Leu | |
| | 370 | | | | | 375 | | | | | 380 | | | | | |
| Asp | Pro | Ser | Asn | Thr | Thr | Pro | Glu | Gly | Val | Ala | Arg | Arg | Tyr | Gly | Gly | |
| 385 | | | | | 390 | | | | | 395 | | | | | 400 | |
| Glu | Trp | Thr | Glu | Asp | Ala | Ala | His | Arg | Ala | Leu | Leu | Ser | Glu | Arg | Leu | |
| | | | | 405 | | | | | 410 | | | | | | 415 | |
| His | Arg | Asn | Leu | Leu | Lys | Arg | Leu | Glu | Gly | Glu | Glu | Lys | Leu | Leu | Trp | |
| | | | 420 | | | | | 425 | | | | | 430 | | | |
| Leu | Tyr | His | Glu | Val | Glu | Lys | Pro | Leu | Ser | Arg | Val | Leu | Ala | His | Met | |
| | | 435 | | | | | 440 | | | | | 445 | | | | |
| Glu | Ala | Thr | Gly | Val | Arg | Arg | Asp | Val | Ala | Tyr | Leu | Gln | Ala | Leu | Ser | |
| | | 450 | | | | 455 | | | | | 460 | | | | | |
| Leu | Glu | Leu | Ala | Glu | Glu | Ile | Arg | Arg | Leu | Glu | Glu | Glu | Val | Phe | Arg | |
| 465 | | | | | 470 | | | | | 475 | | | | | 480 | |
| Leu | Ala | Gly | His | Pro | Phe | Asn | Leu | Asn | Ser | Arg | Asp | Gln | Leu | Glu | Arg | |
| | | | | 485 | | | | | 490 | | | | | 495 | | |
| Val | Leu | Phe | Asp | Glu | Leu | Arg | Leu | Pro | Ala | Leu | Lys | Lys | Thr | Lys | Lys | |
| | | | 500 | | | | | 505 | | | | | 510 | | | |
| Thr | Gly | Lys | Arg | Ser | Thr | Ser | Ala | Ala | Val | Leu | Glu | Ala | Leu | Arg | Glu | |
| | | 515 | | | | | 520 | | | | | 525 | | | | |
| Ala | His | Pro | Ile | Val | Glu | Lys | Ile | Leu | Gln | His | Arg | Glu | Leu | Thr | Lys | |
| | | | | | | 535 | | | | | 540 | | | | | |
| Leu | Lys | Asn | Thr | Tyr | Val | Asp | Pro | Leu | Pro | Ser | Leu | Val | His | Pro | Arg | |
| 545 | | | | | 550 | | | | | 555 | | | | | 560 | |
| Thr | Gly | Arg | Leu | His | Thr | Arg | Phe | Asn | Gln | Thr | Ala | Thr | Ala | Thr | Gly | |
| | | | | 565 | | | | | 570 | | | | | | 575 | |
| Arg | Leu | Ser | Ser | Ser | Asp | Pro | Asn | Leu | Gln | Asn | Ile | Pro | Val | Arg | Thr | |
| | | | 580 | | | | | 585 | | | | | 590 | | | |
| Pro | Leu | Gly | Gln | Arg | Ile | Arg | Arg | Ala | Phe | Val | Ala | Glu | Ala | Gly | Trp | |
| | | 595 | | | | | 600 | | | | | 605 | | | | |
| Ala | Leu | Val | Ala | Leu | Asp | Tyr | Ser | Gln | Ile | Glu | Leu | Arg | Val | Leu | Ala | |
| | | | | | | 615 | | | | | | 620 | | | | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Thr | Thr | Ser | Arg | Gly | Glu | Pro | Val | Gln | Ala | Val | Tyr | Gly | Phe | Ala | Lys |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Ser | Leu | Leu | Lys | Ala | Leu | Arg | Glu | Asp | Gly | Asp | Val | Val | Ile | Val | Val |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| Phe | Asp | Ala | Lys | Ala | Pro | Ser | Phe | Arg | His | Gln | Thr | Tyr | Glu | Ala | Tyr |
| 65 | | | | | 70 | | | | | 75 | | | | | 80 |
| Lys | Ala | Gly | Arg | Ala | Pro | Thr | Pro | Glu | Asp | Phe | Pro | Arg | Gln | Leu | Ala |
| | | | | 85 | | | | | 90 | | | | | 95 | |
| Leu | Ile | Lys | Glu | Met | Val | Asp | Leu | Leu | Gly | Leu | Glu | Arg | Leu | Glu | Val |
| | | | 100 | | | | | 105 | | | | | 110 | | |
| Pro | Gly | Phe | Glu | Ala | Asp | Asp | Val | Leu | Ala | Thr | Leu | Ala | Lys | Lys | Ala |
| | | 115 | | | | | 120 | | | | | 125 | | | |
| Glu | Lys | Glu | Gly | Tyr | Glu | Val | Arg | Ile | Leu | Thr | Ala | Asp | Arg | Asp | Leu |
| | 130 | | | | | 135 | | | | | 140 | | | | |
| Tyr | Gln | Leu | Leu | Ser | Glu | Arg | Ile | Ser | Ile | Leu | His | Pro | Glu | Gly | Tyr |
| 145 | | | | | 150 | | | | | 155 | | | | | 160 |
| Leu | Ile | Thr | Pro | Glu | Trp | Leu | Trp | Glu | Lys | Tyr | Gly | Leu | Lys | Pro | Ser |
| | | | | 165 | | | | | 170 | | | | | 175 | |
| Gln | Trp | Val | Asp | Tyr | Arg | Ala | Leu | Ala | Gly | Asp | Pro | Ser | Asp | Asn | Ile |
| | | | 180 | | | | | 185 | | | | | 190 | | |
| Pro | Gly | Val | Lys | Gly | Ile | Gly | Glu | Lys | Thr | Ala | Ala | Lys | Leu | Ile | Arg |
| | | 195 | | | | | 200 | | | | | 205 | | | |
| Glu | Trp | Gly | Ser | Leu | Glu | Asn | Leu | Leu | Lys | His | Leu | Glu | Gln | Val | Lys |
| | 210 | | | | | 215 | | | | | 220 | | | | |
| Pro | Ala | Ser | Val | Arg | Glu | Lys | Ile | Leu | Ser | His | Met | Glu | Asp | Leu | Lys |
| 225 | | | | | 230 | | | | | 235 | | | | | 240 |
| Leu | Ser | Leu | Glu | Leu | Ser | Arg | Val | His | Thr | Asp | Leu | Leu | Leu | Gln | Val |
| | | | | 245 | | | | | 250 | | | | | 255 | |
| Asp | Phe | Ala | Arg | Arg | Arg | Glu | Pro | Asp | Arg | Glu | Gly | Leu | Lys | Ala | Phe |
| | | | 260 | | | | | 265 | | | | | 270 | | |
| Leu | Glu | Arg | Leu | Glu | Phe | Gly | Ser | Leu | Leu | His | Glu | Phe | Gly | Leu | Leu |
| | | 275 | | | | | 280 | | | | | 285 | | | |
| Glu | Ser | Pro | Val | Ala | Ala | Glu | Glu | Ala | Pro | Trp | Pro | Pro | Pro | Glu | Gly |
| | 290 | | | | | 295 | | | | | 300 | | | | |
| Ala | Phe | Val | Gly | Tyr | Val | Leu | Ser | Arg | Pro | Glu | Pro | Met | Trp | Ala | Glu |
| 305 | | | | | 310 | | | | | 315 | | | | | 320 |
| Leu | Asn | Ala | Leu | Ala | Ala | Ala | Trp | Gly | Gly | Arg | Val | His | Arg | Ala | Ala |
| | | | | 325 | | | | | 330 | | | | | 335 | |

| | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| Asp | Pro | Leu | Ala | Gly | Leu | Lys | Asp | Leu | Lys | Glu | Val | Arg | Gly | Leu | Leu | |
| | | | 340 | | | | | 345 | | | | | | 350 | | |
| Ala | Lys | Asp | Leu | Ala | Val | Leu | Ala | Ser | Arg | Glu | Gly | Leu | Asp | Leu | Val | |
| | | 355 | | | | | 360 | | | | | 365 | | | | |
| Pro | Gly | Asp | Asp | Pro | Met | Leu | Leu | Ala | Tyr | Leu | Leu | Asp | Pro | Ser | Asn | |
| | 370 | | | | | 375 | | | | | 380 | | | | | |
| Thr | Thr | Pro | Glu | Gly | Val | Ala | Arg | Arg | Tyr | Gly | Gly | Glu | Trp | Thr | Glu | |
| 385 | | | | | 390 | | | | | 395 | | | | | 400 | |
| Asp | Ala | Ala | His | Arg | Ala | Leu | Leu | Ser | Glu | Arg | Leu | His | Arg | Asn | Leu | |
| | | | 405 | | | | | | 410 | | | | | 415 | | |
| Leu | Lys | Arg | Leu | Glu | Gly | Glu | Glu | Lys | Leu | Leu | Trp | Leu | Tyr | His | Glu | |
| | | | 420 | | | | | 425 | | | | | 430 | | | |
| Val | Glu | Lys | Pro | Leu | Ser | Arg | Val | Leu | Ala | His | Met | Glu | Ala | Thr | Gly | |
| | | 435 | | | | | 440 | | | | | 445 | | | | |
| Val | Arg | Arg | Asp | Val | Ala | Tyr | Leu | Gln | Ala | Leu | Ser | Leu | Glu | Leu | Ala | |
| | 450 | | | | | 455 | | | | | 460 | | | | | |
| Glu | Glu | Ile | Arg | Arg | Leu | Glu | Glu | Glu | Val | Phe | Arg | Leu | Ala | Gly | His | |
| 465 | | | | | 470 | | | | | 475 | | | | | 480 | |
| Pro | Phe | Asn | Leu | Asn | Ser | Arg | Asp | Gln | Leu | Glu | Arg | Val | Leu | Phe | Asp | |
| | | | 485 | | | | | 490 | | | | | | 495 | | |
| Glu | Leu | Arg | Leu | Pro | Ala | Leu | Lys | Lys | Thr | Lys | Lys | Thr | Gly | Lys | Arg | |
| | | | 500 | | | | | 505 | | | | | 510 | | | |
| Ser | Thr | Ser | Ala | Ala | Val | Leu | Glu | Ala | Leu | Arg | Glu | Ala | His | Pro | Ile | |
| | | 515 | | | | | 520 | | | | | 525 | | | | |
| Val | Glu | Lys | Ile | Leu | Gln | His | Arg | Glu | Leu | Thr | Lys | Leu | Lys | Asn | Thr | |
| | 530 | | | | | 535 | | | | | 540 | | | | | |
| Tyr | Val | Asp | Pro | Leu | Pro | Ser | Leu | Val | His | Pro | Arg | Thr | Gly | Arg | Leu | |
| 545 | | | | | 550 | | | | | 555 | | | | | 560 | |
| His | Thr | Arg | Phe | Asn | Gln | Thr | Ala | Thr | Ala | Thr | Gly | Arg | Leu | Ser | Ser | |
| | | | 565 | | | | | 570 | | | | | | 575 | | |
| Ser | Asp | Pro | Asn | Leu | Gln | Asn | Ile | Pro | Val | Arg | Thr | Pro | Leu | Gly | Gln | |
| | | | 580 | | | | | 585 | | | | | 590 | | | |
| Arg | Ile | Arg | Arg | Ala | Phe | Val | Ala | Glu | Ala | Gly | Trp | Ala | Leu | Val | Ala | |
| | | 595 | | | | | 600 | | | | | 605 | | | | |
| Leu | Asp | Tyr | Ser | Gln | Ile | Glu | Leu | Arg | Val | Leu | Ala | His | Leu | Ser | Gly | |
| | 610 | | | | | 615 | | | | | 620 | | | | | |
| Asp | Glu | Asn | Leu | Ile | Arg | Val | Phe | Gln | Glu | Gly | Lys | Asp | Ile | His | Thr | |
| 625 | | | | | 630 | | | | | 635 | | | | | 640 | |

Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu Ala Val Asp Pro
645 650 655

Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly Val Leu Tyr Gly
660 665 670

Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu Glu
675 680 685

Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val Arg
690 695 700

Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys Arg Gly Tyr Val
705 710 715 720

Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn Ala Arg
725 730 735

Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro
740 745 750

Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu
755 760 765

Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu Leu Gln Val Ala
770 775 780

Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala Glu Glu Val Ala
785 790 795 800

Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro Leu Ala Val Pro
805 810 815

Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala Lys Gly
820 825 830

His His His His His His
835

<210> 67
<211> 842
<212> PRT
<213> Artificial

<220>
<223> Synthetic

<400> 67

Met Asn Ser Thr Pro Leu Phe Asp Leu Glu Glu Pro Pro Lys Arg Val
1 5 10 15
Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Tyr Ala Leu
20 25 30

Ser Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Met Val Tyr Gly Phe
35 40 45

Ala Arg Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Gln Ala Val Val
50 55 60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu
65 70 75 80

Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
85 90 95

Leu Ala Leu Val Lys Arg Leu Val Asp Leu Leu Gly Leu Val Arg Leu
100 105 110

Glu Ala Pro Gly Tyr Glu Ala Asp Asp Val Leu Gly Thr Leu Ala Lys
115 120 125

Lys Ala Glu Arg Glu Gly Met Glu Val Arg Ile Leu Thr Gly Asp Arg
130 135 140

Asp Phe Phe Gln Leu Leu Ser Glu Lys Val Ser Val Leu Leu Pro Asp
145 150 155 160

Gly Thr Leu Val Thr Pro Lys Asp Val Gln Glu Lys Tyr Gly Val Pro
165 170 175

Pro Glu Arg Trp Val Asp Phe Arg Ala Leu Thr Gly Asp Arg Ser Asp
180 185 190

Asn Ile Pro Gly Val Ala Gly Ile Gly Glu Lys Thr Ala Leu Arg Leu
195 200 205

Leu Ala Glu Trp Gly Ser Val Glu Asn Leu Leu Lys Asn Leu Asp Arg
210 215 220

Val Lys Pro Asp Ser Leu Arg Arg Lys Ile Glu Ala His Leu Glu Asp
225 230 235 240

Leu His Leu Ser Leu Asp Leu Ala Arg Ile Arg Thr Asp Leu Pro Leu
245 250 255

Glu Val Asp Phe Lys Ala Leu Arg Arg Arg Thr Pro Asp Leu Glu Gly
260 265 270

Leu Arg Ala Phe Leu Glu Glu Leu Glu Phe Gly Ser Leu Leu His Glu
275 280 285

Phe Gly Leu Leu Gly Gly Glu Lys Pro Arg Glu Glu Ala Pro Trp Pro
290 295 300

Pro Pro Glu Gly Ala Phe Val Gly Phe Leu Leu Ser Arg Lys Glu Pro
305 310 315 320

Met Trp Ala Glu Leu Leu Ala Leu Ala Ala Ala Ser Gly Gly Arg Val
325 330 335

His Arg Ala Pro Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala
340 345 350

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Arg | Gly | Leu | Leu | Ala | Lys | Asp | Leu | Ser | Val | Leu | Ala | Leu | Arg | Glu | Gly |
| | | 355 | | | | | 360 | | | | | 365 | | | |
| Leu | Gly | Leu | Pro | Pro | Gly | Asp | Asp | Pro | Met | Leu | Leu | Ala | Tyr | Leu | Leu |
| | 370 | | | | | 375 | | | | | 380 | | | | |
| Asp | Pro | Ser | Asn | Thr | Thr | Pro | Glu | Gly | Val | Ala | Arg | Arg | Tyr | Gly | Gly |
| 385 | | | | | 390 | | | | | 395 | | | | | 400 |
| Glu | Trp | Thr | Glu | Glu | Ala | Gly | Glu | Arg | Ala | Ala | Leu | Ser | Glu | Arg | Leu |
| | | | | 405 | | | | | 410 | | | | | 415 | |
| Phe | Ala | Asn | Leu | Leu | Lys | Arg | Leu | Glu | Gly | Glu | Glu | Arg | Leu | Leu | Trp |
| | | | 420 | | | | | 425 | | | | | 430 | | |
| Leu | Tyr | Arg | Glu | Val | Glu | Arg | Pro | Leu | Ser | Ala | Val | Leu | Ala | His | Met |
| | | 435 | | | | | 440 | | | | | 445 | | | |
| Glu | Ala | Thr | Gly | Val | Arg | Leu | Asp | Val | Ala | Tyr | Leu | Arg | Ala | Leu | Ser |
| | 450 | | | | | 455 | | | | | 460 | | | | |
| Leu | Glu | Val | Ala | Glu | Glu | Ile | Ala | Arg | Leu | Glu | Ala | Glu | Val | Phe | Arg |
| 465 | | | | | 470 | | | | | 475 | | | | | 480 |
| Leu | Ala | Gly | His | Pro | Phe | Asn | Leu | Asn | Ser | Arg | Asp | Gln | Leu | Glu | Arg |
| | | | | 485 | | | | | 490 | | | | | 495 | |
| Val | Leu | Phe | Asp | Glu | Leu | Gly | Leu | Pro | Ala | Ile | Lys | Lys | Thr | Gln | Lys |
| | | | 500 | | | | | 505 | | | | | 510 | | |
| Thr | Gly | Lys | Arg | Ser | Thr | Ser | Ala | Ala | Val | Leu | Glu | Ala | Leu | Arg | Glu |
| | | 515 | | | | | 520 | | | | | 525 | | | |
| Ala | His | Pro | Ile | Val | Glu | Lys | Ile | Leu | Gln | Tyr | Arg | Glu | Leu | Thr | Lys |
| | 530 | | | | | 535 | | | | | 540 | | | | |
| Leu | Lys | Ser | Thr | Tyr | Ile | Asp | Pro | Leu | Pro | Asp | Leu | Ile | His | Pro | Arg |
| 545 | | | | | 550 | | | | | 555 | | | | | 560 |
| Thr | Gly | Arg | Leu | His | Thr | Arg | Phe | Asn | Gln | Thr | Ala | Thr | Ala | Thr | Gly |
| | | | | 565 | | | | | 570 | | | | | 575 | |
| Arg | Leu | Ser | Ser | Ser | Asp | Pro | Asn | Leu | Gln | Asn | Ile | Pro | Val | Arg | Thr |
| | | | 580 | | | | | 585 | | | | | 590 | | |
| Pro | Leu | Gly | Gln | Arg | Ile | Arg | Arg | Ala | Phe | Ile | Ala | Glu | Glu | Gly | Trp |
| | | 595 | | | | | 600 | | | | | 605 | | | |
| Leu | Leu | Val | Ala | Leu | Asp | Tyr | Ser | Gln | Ile | Glu | Leu | Arg | Val | Leu | Ala |
| | 610 | | | | | 615 | | | | | 620 | | | | |
| His | Leu | Ser | Gly | Asp | Glu | Asn | Leu | Ile | Arg | Val | Phe | Gln | Glu | Gly | Arg |
| 625 | | | | | 630 | | | | | 635 | | | | | 640 |
| Asp | Ile | His | Thr | Glu | Thr | Ala | Ser | Trp | Met | Phe | Gly | Val | Pro | Arg | Glu |
| | | | | 645 | | | | | 650 | | | | | 655 | |

Ala Val Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly
660 665 670

Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile
675 680 685

Pro Tyr Glu Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe
690 695 700

Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg
705 710 715 720

Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp
725 730 735

Leu Glu Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala
740 745 750

Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala
755 760 765

Met Val Lys Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu
770 775 780

Leu Gln Val Ala Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala
785 790 795 800

Glu Ala Val Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro
805 810 815

Leu Ala Val Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu
820 825 830

Ser Ala Lys Glu His His His His His His
835 840

<210> 68
<211> 838
<212> PRT
<213> Artificial

<220>
<223> Synthetic

<400> 68

Met Asn Ser Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu Val
1 5 10 15

Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu Lys Gly Leu
20 25 30

Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala Lys
35 40 45

Ser Leu Leu Lys Ala Leu Arg Glu Asp Gly Asp Val Val Ile Val Val

| | | |
|---|-----|-------------|
| 50 | 55 | 60 |
| Phe Asp Ala Lys Ala Pro Ser Phe Arg His Gln Thr Tyr Glu Ala Tyr | | |
| 65 | 70 | 75 80 |
| Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu Ala | | |
| | 85 | 90 95 |
| Leu Ile Lys Glu Met Val Asp Leu Leu Gly Leu Glu Arg Leu Glu Val | | |
| | 100 | 105 110 |
| Pro Gly Phe Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys Lys Ala | | |
| | 115 | 120 125 |
| Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Arg Asp Leu | | |
| | 130 | 135 140 |
| Tyr Gln Leu Leu Ser Glu Arg Ile Ser Ile Leu His Pro Glu Gly Tyr | | |
| | 145 | 150 155 160 |
| Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly Leu Lys Pro Ser | | |
| | 165 | 170 175 |
| Gln Trp Val Asp Tyr Arg Ala Leu Ala Gly Asp Pro Ser Asp Asn Ile | | |
| | 180 | 185 190 |
| Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Ala Lys Leu Ile Arg | | |
| | 195 | 200 205 |
| Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys His Leu Glu Gln Val Lys | | |
| | 210 | 215 220 |
| Pro Ala Ser Val Arg Glu Lys Ile Leu Ser His Met Glu Asp Leu Lys | | |
| | 225 | 230 235 240 |
| Leu Ser Leu Glu Leu Ser Arg Val His Thr Asp Leu Leu Leu Gln Val | | |
| | 245 | 250 255 |
| Asp Phe Ala Arg Arg Arg Glu Pro Asp Arg Glu Gly Leu Lys Ala Phe | | |
| | 260 | 265 270 |
| Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu Leu | | |
| | 275 | 280 285 |
| Glu Ser Pro Val Ala Ala Glu Glu Ala Pro Trp Pro Pro Pro Glu Gly | | |
| | 290 | 295 300 |
| Ala Phe Val Gly Tyr Val Leu Ser Arg Pro Glu Pro Met Trp Ala Glu | | |
| | 305 | 310 315 320 |
| Leu Asn Ala Leu Ala Ala Ala Trp Gly Gly Arg Val His Arg Ala Pro | | |
| | 325 | 330 335 |
| Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu Leu | | |
| | 340 | 345 350 |
| Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu Pro | | |

| | | |
|---|-----|-----|
| 355 | 360 | 365 |
| Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser Asn | | |
| 370 | 375 | 380 |
| Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr Glu | | |
| 385 | 390 | 395 |
| Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn Leu | | |
| | 405 | 410 |
| Leu Lys Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg Glu | | |
| | 420 | 425 |
| Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr Gly | | |
| | 435 | 440 |
| Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val Ala | | |
| | 450 | 455 |
| Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly His | | |
| 465 | 470 | 475 |
| Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe Asp | | |
| | 485 | 490 |
| Glu Leu Gly Leu Pro Ala Ile Lys Lys Thr Gln Lys Thr Gly Lys Arg | | |
| | 500 | 505 |
| Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro Ile | | |
| | 515 | 520 |
| Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser Thr | | |
| | 530 | 535 |
| Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg Leu | | |
| 545 | 550 | 555 |
| His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser Ser | | |
| | 565 | 570 |
| Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly Gln | | |
| | 580 | 585 |
| Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val Ala | | |
| | 595 | 600 |
| Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser Gly | | |
| | 610 | 615 |
| Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His Thr | | |
| 625 | 630 | 635 |
| Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp Pro | | |
| | 645 | 650 |
| Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr Gly | | |

660 665 670

Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu Glu
675 680 685

Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val Arg
690 695 700

Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr Val
705 710 715 720

Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala Arg
725 730 735

Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro
740 745 750

Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu
755 760 765

Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val Ala
770 775 780

Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val Ala
785 790 795 800

Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val Pro
805 810 815

Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys Glu
820 825 830

His His His His His His
835

<210> 69
<211> 2517
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 69
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
ccggtgcagg cggctctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggccgggc cccaacgccg gaggactttc cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggctggcg cgctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatactc 420

| | | | | | | |
|------------|------------|------------|------------|------------|------------|------|
| accgccgaca | aagaccttta | ccagctcctt | tccgaccgca | tccacgtcct | ccaccccag | 480 |
| gggtacctca | tcaccccggc | ctggcttttg | gaaaagtacg | gcctgaggcc | cgaccagtgg | 540 |
| gccgactacc | gggccctgac | cggggacgag | tccgacaacc | ttcccggggg | caagggcatc | 600 |
| ggggagaaga | cggcgaggaa | gcttctggag | gagtggggga | gcctggaagc | cctcctcaag | 660 |
| aacctggacc | ggctgaagcc | cgccatccgg | gagaagatcc | tggcccat | ggacgatctg | 720 |
| aagctctcct | gggacctggc | caaggtgcgc | accgacctgc | ccctggaggt | ggacttcgcc | 780 |
| aaaaggcggg | agcccgaccg | ggagaggctt | agggcctttc | tggagaggct | tgagtttggc | 840 |
| agcctcctcc | acgagttcgg | ccttctggaa | agccccaagg | ccctggagga | ggccccctgg | 900 |
| ccccgcgcgg | aaggggcctt | cgtgggcctt | gtgctttccc | gcaaggagcc | catgtggggc | 960 |
| gatctttctg | ccctggccgc | cgccaggggc | ggccgcgtgc | accgggcagc | agacccttg | 1020 |
| gcggggctaa | aggacctcaa | ggaggtccgg | ggcctcctcg | ccaaggacct | cgccgtcttg | 1080 |
| gcctcgaggg | aggggctaga | cctcgtgccc | ggggacgacc | ccatgctcct | cgctacctc | 1140 |
| ctggaccctt | cgaacaccac | ccccgagggg | gtggcgcggc | gctacggggg | ggagtggacg | 1200 |
| gaggacgccg | cccaccgggc | cctcctctcg | gagaggctcc | atcggaacct | ccttaagcgc | 1260 |
| ctcgaggggg | aggagaagct | cctttggctc | taccacgagg | tggaaaagcc | cctctcccgg | 1320 |
| gtcctggccc | atatggaggc | caccggggta | cggcgggacg | tggcctacct | tcaggccctt | 1380 |
| tccctggagc | ttgcggagga | gatccgccgc | ctcgaggagg | aggtcttccg | cttggcgggc | 1440 |
| cacccttca | acctcaactc | ccgggaccag | ctggaaaggg | tgctctttga | cgagcttagg | 1500 |
| cttcccgcct | tggggaagac | gcaaaagaca | ggcaagcgct | ccaccagcgc | cgcggtgctg | 1560 |
| gaggccctac | gggaggccca | ccccatcgtg | gagaagatcc | tccagcaccg | ggagctcacc | 1620 |
| aagctcaaga | acacctacgt | ggacccccct | ccaagcctcg | tccacccgag | gacgggccgc | 1680 |
| ctccacaccc | gcttcaacca | gacggccacg | gccacgggga | ggcttagtag | ctccgacccc | 1740 |
| aacctgcaga | acatccccgt | ccgcaccccc | ttgggccaga | ggatccgcgc | ggccttcgtg | 1800 |
| gccgaggcgg | gttgggcggt | ggtggccctg | gactatagcc | agatagagct | ccgcgtcctc | 1860 |
| gcccacctct | ccggggacga | aaacctgatc | agggctcttc | aggaggggaa | ggacatccac | 1920 |
| accagaccg | caagctggat | gttcggcgtc | cccccgagg | ccgtggaccc | cctgatgcgc | 1980 |
| cgggcggcca | agacggtgaa | cttcggcgtc | ctctacggca | tgtccgcca | taggctctcc | 2040 |
| caggagcttg | ccatccccta | cgaggaggcg | gtggccttta | tagagcgcta | cttccaaagc | 2100 |

| | |
|--|------|
| ttccccaagg tgcgggcctg gatagaaaag accctggagg aggggaggaa gcggggctac | 2160 |
| gtggaaaccc tcttcggaag aaggcgctac gtgcccagacc tcaacgcccg ggtgaagagc | 2220 |
| gtcagggagg ccgcgagcg catggccttc aacatgcccg tccagggcac cgccgcccac | 2280 |
| ctcatgaagc tcgccatggt gaagctcttc ccccgccctcc gggagatggg ggcccgcctg | 2340 |
| ctcctccagg tccacaacga gtcctcctg gagggccccc aagcgccggc cgaggagggtg | 2400 |
| gcggtcttgg ccaaggaggc catggagaag gcctatcccc tcgccgtgcc cctggagggtg | 2460 |
| gaggtgggga tgggggagga ctggctttcc gccaaagggtc accaccacca ccaccac | 2517 |

<210> 70
 <211> 2526
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

| | |
|--|------|
| <400> 70 | |
| atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac | 60 |
| ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc | 120 |
| gaaccggtgc aggcggtcta cggtctcgcc aagagcctcc tcaaggccct gaaggaggac | 180 |
| gggtacaagg ccgtcttcgt ggtctttgac gccaaaggccc cctccttccg ccacgaggcc | 240 |
| tacgaggcct acaaggcggg gagggccccc acccccaggg acttcccccg gcagctcgcc | 300 |
| ctcatcaagg agctggtgga cctcctgggg ttaccgcc tcgagggtccc cggctacgag | 360 |
| gcggacgacg ttctcgccac cctggccaag aaggcggaag aggaggggta cgagggtgcg | 420 |
| atcctcaccg ccgaccgga cctctaccaa ctcgctcctc accgctcgc cgtcctccac | 480 |
| cccaggggcc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag | 540 |
| cagtgggtgg acttcgcgc cctcgtgggg gaccctcctc acaacctccc cggggtcaag | 600 |
| ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc | 660 |
| ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg | 720 |
| gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctccc cctggagggtg | 780 |
| gacctcgccc aggggcggga gcccagccgg gaggggctta gggccttcct ggagaggctg | 840 |
| gagttcggca gcctcctcca cgagttcggc ctcttgagg ccccgcccc cctggaggag | 900 |
| gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctccc ccccgagccc | 960 |
| atgtgggcgg agcttaaagc cctggccgcc tgcaggggag gccgcgtcca ccgggcccc | 1020 |

| | | | | | | |
|------------|-------------|-------------|------------|------------|------------|------|
| gagccttata | aagccctcag | ggacctgaag | gaggcgcggg | ggcttctcgc | caaagacctg | 1080 |
| agcgttcttg | ccctgagggg | aggccttggc | ctcccgcccg | gcgacgaccc | catgctcctc | 1140 |
| gcctacctcc | tggaccttcc | gaacaccacc | cccgaggggg | tggcccggcg | ctacggcggg | 1200 |
| gagtggacgg | aggaggcggg | ggagcgggcc | gccctttccg | agaggctctt | cgccaacctg | 1260 |
| tgggggaggg | ttgaggggga | ggagaggctc | ctttggcttt | accgggaggt | ggagaggccc | 1320 |
| ctttccgctg | tcctggccca | tatggaggcc | acgggggtgc | gcctggacgt | ggcctatctc | 1380 |
| agggccttgt | ccctggaggt | ggccgaggag | atcgcccgcc | tcgaggccga | ggtcttcgcg | 1440 |
| ctggccggcc | accccttcaa | cctcaactcc | cgggaccagc | tggaaagggt | cctctttgac | 1500 |
| gagctagggc | ttcccgccat | cggcaagacg | gagaagaccg | gcaagcgctc | caccagcgcc | 1560 |
| gccgtcctgg | aggccctccg | cgaggcccac | cccacgtgg | agaagatcct | gcagtaccgg | 1620 |
| gagctcacca | agctgaagag | cacctacatt | gaccccttgc | cggacctcat | ccaccccagg | 1680 |
| acgggcccgc | tccacacccg | cttcaaccag | acggccacgg | ccacgggcag | gctaagtagc | 1740 |
| tccgatccca | acctccagaa | catccccgtc | cgcaccccgc | ttgggcagag | gatccgcccg | 1800 |
| gccttcatcg | ccgaggaggg | gtggctattg | gtggccctgg | actatagcca | gatagagctc | 1860 |
| aggggtgctg | cccacctctc | cggcgacgag | aacctgatcc | gggtcttcca | ggaggggcgg | 1920 |
| gacatccaca | cggagaccgc | cagctggatg | ttcggcgctc | cccgggaggc | cgtggacccc | 1980 |
| ctgatgcgcc | gggcggccaa | gaccatcaac | ttcggggctc | tctacggcat | gtcggcccac | 2040 |
| cgcctctccc | aggagctagc | catccccttac | gaggaggccc | aggccttcat | tgagcgctac | 2100 |
| tttcagagct | tccccaagggt | gcgggcctgg | attgagaaga | ccctggagga | gggcaggagg | 2160 |
| cgggggtacg | tggagaccct | cttcggcccg | cgcgctacg | tgccagacct | agaggcccgg | 2220 |
| gtgaagagcg | tgcgggaggc | ggccgagcgc | atggccttca | acatgcccgt | ccagggcacc | 2280 |
| gccgccgacc | tcataaagct | ggctatggtg | aagctcttcc | ccaggctgga | ggaaatgggg | 2340 |
| gccaggatgc | tccttcagggt | ccacaacgag | ctggtcctcg | aggccccaaa | agagagggcg | 2400 |
| gaggccgtgg | cccggctggc | caaggaggtc | atggaggggg | tgtatccoct | ggccgtgccc | 2460 |
| ctggaggtgg | aggtggggat | aggggaggac | tggctctccg | ccaaggagca | ccaccaccac | 2520 |
| caccac | | | | | | 2526 |

```
<210> 71
<211> 2517
<212> DNA
<213> Artificial
```


| | | | | | | |
|-------------|------------|-------------|-------------|------------|------------|------|
| gaggccctac | gggaggccca | ccccatcgtg | gagaagatcc | tccagcaccg | ggagctcacc | 1620 |
| aagctcaaga | acacctacgt | ggacccccctc | ccaagcctcg | tccacccgag | gacgggccgc | 1680 |
| ctccacaccc | gcttcaacca | gacggccacg | gccacgggga | ggcttagtag | ctccgacccc | 1740 |
| aacctgcaga | acatccccgt | ccgcaccccc | ttggggccaga | ggatccgcgc | ggccttcac | 1800 |
| gccgaggagg | ggtggctatt | ggtggccctg | gactatagcc | agatagagct | cagggtgctg | 1860 |
| gccacctct | ccggcgacga | gaacctgatc | cgggtcttcc | aggaggggcg | ggacatccac | 1920 |
| acggagaccg | ccagctggat | gttcggcgtc | ccccgggagg | ccgtggaccc | cctgatgcgc | 1980 |
| cgggcgggcca | agaccatcaa | cttcggggtc | ctctacggca | tgtcggccca | ccgcctctcc | 2040 |
| caggagctag | ccatccctta | cgaggaggcc | caggcettca | ttgagcgcta | ctttcagagc | 2100 |
| ttccccaagg | tgcgggcctg | gattgagaag | accctggagg | agggcaggag | gcgggggtac | 2160 |
| gtggagaccc | tcttcggccg | ccgccgctac | gtgccagacc | tagaggcccg | ggtgaagagc | 2220 |
| gtgcgggagg | cggccgagcg | catggccttc | aacatgcccg | tccagggcac | cgccgccgac | 2280 |
| ctcatgaagc | tggctatggt | gaagctcttc | cccaggctgg | aggaaatggg | ggccaggatg | 2340 |
| ctccttcagg | tccacaacga | gctggtcctc | gaggccccaa | aagagagggc | ggaggccgtg | 2400 |
| gcccggtgg | ccaaggaggt | catggagggg | gtgtatcccc | tggccgtgcc | cctggaggtg | 2460 |
| gagggtgggga | taggggagga | ctggctctcc | gccaaggagc | accaccacca | ccaccac | 2517 |

```
<210> 72
<211> 2517
<212> DNA
<213> Artificial
```

<220>
<223> Synthetic

| | | | | | | | |
|------------|------------|------------|------------|------------|-------------|--|-----|
| <400> | 72 | | | | | | |
| atgaattcgg | ggatgctgcc | cctctttgag | cccaagggcc | gggtcctcct | ggtggacggc | | 60 |
| caccacctgg | cctaccgcac | cttccacgcc | ctgaagggcc | tcaccaccag | ccgggggggag | | 120 |
| ccggtgcagg | cggtctacgg | cttcgccaag | agcctcctca | aggccctcaa | ggaggacggg | | 180 |
| gacgcggtga | tcgtggtctt | tgacgccaag | gccccctcct | tccgccacga | ggcctacggg | | 240 |
| gggtacaagg | cgggccgggc | ccccacgccg | gaggactttc | cccggcaact | cgccctcatc | | 300 |
| aaggagctgg | tggacctcct | ggggctggcg | cgctctgagg | tcccgggcta | cgaggcggac | | 360 |
| gacgtcctgg | ccagcctggc | caagaaggcg | gaaaaggagg | gctacgaggt | ccgcctcctc | | 420 |
| accgccgaca | aagaccttta | ccagctcctt | tccgaccgca | tccacgtcct | ccacccccgag | | 480 |

ggggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
 gccgactacc gggccctgac cggggacgag tccgacaacc ttcccgggggt caagggcatc 600
 ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
 aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccatat ggacgatctg 720
 aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggagggt ggacttcgcc 780
 aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
 agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
 cccccgcggg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtggggc 960
 gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
 aaagccctca gggacctgaa ggaggcgagg gggcttctcg ccaaagacct gagcgttctg 1080
 gccctgaggg aaggccttgg cctcccgccc ggcgacgacc ccatgctcct cgcctacctc 1140
 ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200
 gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gtgggggagg 1260
 cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct 1320
 gtcttgccc atatggaggc cacgggggtg cgcttgacg tggcctatct cagggccttg 1380
 tccctggagg tggccgagga gatcgccgcg ctcgaggcgg aggtcttccg cctggccggc 1440
 cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
 cttcccgcca tcggcaagac ggagaagacc ggcaagcgt ccaccagcgc cgccgtcctg 1560
 gaggcctcc gcgaggcca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
 aagctgaaga gcacctacat tgaccccttg ccggacctca tccaccccag gacgggcccgc 1680
 ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
 aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgcgg ggccttcgtg 1800
 gccgaggcgg gttgggcgtt ggtggccctg gactatagcc agatagagct ccgcgtcctc 1860
 gccacctct ccggggacga aaacctgate agggctcttc aggaggggaa ggacatccac 1920
 acccagaccg caagctggat gttcggcgtc ccccggagg ccgtggaccc cctgatgcgc 1980
 cgggcccga agacggtgaa cttcggcgtc ctctacggca tgtccgcca taggctctcc 2040
 caggagcttg ccatccccta cgaggaggcg gtggccttta tagagcgcta cttccaaagc 2100
 ttccccaagg tgcgggcctg gatagaaaag accctggagg aggggaggaa gcggggctac 2160

gtggaaaccc tcttcggaag aaggcgctac gtgcccgcacc tcaacgcccgc ggtgaagagc 2220
 gtcagggagg ccgcggagcg catggccttc aacatgcccgc tccagggcac cgccgcccgc 2280
 ctcatgaagc tcgcatggt gaagctcttc ccccgccctcc gggagatggg ggcccgcagc 2340
 ctctccagg tccacaacga gctcctcctg gagggccccc aagcgccggc cgaggaggcg 2400
 gcggctttgg ccaaggaggc catggagaag gcctatcccc tcgcccgtgc cctggaggcg 2460
 gaggtgggga tgggggagga ctggctttcc gccaaagggtc accaccacca ccaccac 2517

<210> 73
 <211> 2517
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 73
 atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
 caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
 ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
 gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggccctacggg 240
 ggggtacaagg cgggcccgggc cccacgcgcg gaggactttc cccggcaact cgccctcctc 300
 aaggagctgg tggacctcct ggggctggcg cgctcgagg tcccgggcta cgaggcggac 360
 gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420
 accgcccaga aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccag 480
 ggggtacctca tcaccccggc ctggctttgg gaaaagtacg gcctgaggcc cgaccagtgg 540
 gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggc caagggcatc 600
 ggggagaaga cggcgaggaa gcttctggag gagggggga gcctggaagc cctcctcaag 660
 aacctggacc ggetgaagcc cgccatccgc gagaagatcc tggcccatat ggacgatctg 720
 aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggc ggacttcgcc 780
 aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
 agcctcctcc acgagttcgc ccttctggaa agcccgaagg ccctggagga ggccccctgg 900
 ccccgccgcg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggccc 960
 gatcttctgg ccctggccgc cgccaggggc ggccgcgtgc accgggcagc agacccttg 1020
 gcgggggctaa aggacctcaa ggaggctcgc ggccctcctc ccaaggacct cgccgtcttg 1080

gcctcgaggg aggggctaga cctcgtgccc ggggacgacc ccatgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccagaggg gtggcgcggc gctacggggg ggagtggacg 1200
gaggacgccg cccaccgggc cctcctctcg gagaggctcc atcggaacct ccttaagcgc 1260
ctcgaggggg aggagaagct cctttggctc taccacgagg tggaaaagcc cctctcccgg 1320
gtcctggccc atatggaggc cacgggggtg cgcttggacg tggcctatct cagggccttg 1380
tccctggagg tggccgagga gatcgccgcg ctcgagggcg aggtcttccg cctggccggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
cttcccgcga tcggcaagac ggagaagacc ggcaagcgct ccaccagcgc cgccgtcctg 1560
gaggccctcc gcgaggccca ccccatcggt gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgacccttg ccggacctca tccacccag gacgggcccgc 1680
ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggcccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
gcccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac 1920
acggagaccg ccagctggat gttcggcgtc ccccgaggag ccgtggaccc cctgatgcgc 1980
cgggcggcca agaccatcaa cttcggggtc ctctacggca tgcgggcca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
ctcatgaagc tggctatggt gaagctcttc ccaggtctgg aggaaatggg ggccaggatg 2340
ctccttcagg tccacaacga gctggctctc gagggcccaa aagagagggc ggaggccgtg 2400
gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggag accaccacca ccaccac 2517

<210> 74
<211> 2517
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 74
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggcccggg cccacgccc gaggaactttc cccggcaact cgcctcatc 300
aaggagctgg tggacctcct ggggctggcg cgctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccgcc ctggccttgg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg 720
aagctctcct gggacctggc caagggtgcg accgacctgc ccctggagggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga gggccctcgg 900
ccccgcgagg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggccc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
aaagccctca gggacctgaa ggaggcgagg gggcttctcg ccaaagacct gagcgttctg 1080
gccctgaggg aaggccttgg cctcccgcc ggcgacgacc ccatgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccagggg gtggcccggc gctacggcgg ggagtggacg 1200
gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gtggggggagg 1260
cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct 1320
gtcctggccc atatggaggc caccggggta cggcgggacg tggcctacct tcaggccctt 1380
tccttgagc ttgcggagga gatccgcgc ctcgaggagg aggtcttccg cttggcgggc 1440
cacccttca acctcaactc cggggaccag ctggaaaggg tgctctttga cgagcttagg 1500
cttccgcct tggggaagac gcaaaagaca ggcaagcgt ccaccagcg cgcggtgctg 1560
gaggccctac gggaggcca cccatcgtg gagaagatcc tccagcaccg ggagctcacc 1620
aagctcaaga acacctacgt ggacccctc ccaagcctcg tccacccgag gacgggcccgc 1680

| | |
|--|------|
| ctccacaccc gcttcaacca gacggccacg gccacgggga ggcttagtag ctccgacccc | 1740 |
| aacctgcaga acatccccgt ccgcaccccc ttgggccaga ggatccgccg ggcccttcac | 1800 |
| gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct caggggtgctg | 1860 |
| gcccacctct ccggcgacga gaacctgac cggtctctcc aggaggggag ggacatccac | 1920 |
| acggagaccg ccagctggat gttcggcgtc ccccgaggag ccgtggaccc cctgatgcgc | 1980 |
| cgggcgccca agaccatcaa cttcggggtc ctctacggca tgtcggccca ccgcctctcc | 2040 |
| caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc | 2100 |
| ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac | 2160 |
| gtggagaccc tcttcggccg ccgcgcgtac gtgccagacc tagaggcccg ggtgaagagc | 2220 |
| gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac | 2280 |
| ctcatgaagc tggctatggt gaagctcttc ccaggctgg aggaaatggg ggccaggatg | 2340 |
| ctccttcagg tccacaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg | 2400 |
| gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggagggtg | 2460 |
| gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac | 2517 |

<210> 75
 <211> 2517
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

| | |
|---|-----|
| <400> 75 | |
| atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc | 60 |
| caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag | 120 |
| ccggtgcagg cggctctacg ctctgccaag agcctcctca aggccctcaa ggaggacggg | 180 |
| gacgcggtga tcgtggtctt tgacgccaaag gccccctcct tccgccaaga ggccctacggg | 240 |
| gggtacaagg cgggccgggc ccccaacgcc gaggaacttc cccggcaact cgccctcatc | 300 |
| aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac | 360 |
| gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc | 420 |
| accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag | 480 |
| gggtacctca tcaccccgcc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg | 540 |

gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggt caagggcatc 600
 ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
 aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggccacat ggacgatctg 720
 aagctctcct gggacctggc caaggtgcmc accgacctgc ccctggaggt ggacttcgcc 780
 aaaaggcggg agcccgaccg ggagaggctt agggccttcc tggagaggct tgagtttgcc 840
 agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga gggccctgg 900
 ccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc 960
 gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
 aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg 1080
 gccctgaggg aaggccttgg cctccgccc ggcgacgacc ccatgctcct cgcctacctc 1140
 ctggaccctt cgaacaccac ccccgagggg gtggcgccgc gctacggggg ggagtggacg 1200
 gaggacgccg ccacccgggc cctcctctcg gagaggctcc atcggaacct ccttaagcgc 1260
 ctcgaggggg aggagaagct cctttggctc taccacgagg tggaaaagcc cctctcccgg 1320
 gtccctggccc atatggaggc caccggggta cggcgggacg tggcctacct tcaggccctt 1380
 tccctggagc ttgaggagga gatccgccgc ctcgaggagg aggtcttccg cttggcgggc 1440
 cacccttca acctcaactc ccgggaccag ctggaaaggg tgctctttga cgagcttagg 1500
 cttcccgct tggggaagac gaaaagaca ggcaagcgt ccaccagcgc cgcggtgctg 1560
 gaggccctac gggaggccca cccatcgtg gagaagatcc tccagcaccg ggagctcacc 1620
 aagctcaaga acacctacgt ggacccctc ccaagcctcg tccacccgag gacgggcccgc 1680
 ctccacaccc gcttcaacca gacggccacg gccacgggga ggcttagtag ctccgacccc 1740
 aacctgcaga acatccccgt ccgcaccccc ttgggcccaga ggatccgccg ggccttcac 1800
 gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
 gccacctct ccggcgacga gaacctgac cgggtcttcc aggagggcg ggacatccac 1920
 acggagaccg ccagctggat gtccggcgtc ccccgggagg ccgtggaccc cctgatgcgc 1980
 cgggcggcca agaccatcaa ctccggggtc ctctacggca tgcggccca ccgcctctcc 2040
 caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 tcccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
 gtggagaccc tcttcggccg ccgcgcctac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280

ctcatgaagc tggctatggt gaagctcttc cccaggctgg aggaaatggg ggccaggatg 2340
ctccttcagg tccacaacga gctggtcctc gaggcccaa aagagagggc ggaggccgtg 2400
gcccggtcgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 76
<211> 2517
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 76
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggctcctct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
ccggtgcagg cggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggcccggc cccacgcgcg gaggactttc cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccag 480
gggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggt caagggcac 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggccacat ggacgatctg 720
aagctctcct gggacctggc caaggtgcmc accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
ccccgcgcgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtgc accgggcagc agacccttg 1020
gcggggctaa aggacctcaa ggaggtccgg ggccctcctc ccaaggacct cgccgtcttg 1080
gcctcgaggg aggggctaga cctcgtgccc ggggacgacc ccattgctcct cgccctac 1140

ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200
gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gtgggggagg 1260
cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct 1320
gtcctggccc atatggaggc cacgggggtg cgcttgacg tggcctatct cagggccttg 1380
tccctggagg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
cttcccgcga tcggcaagac ggagaagacc ggcaagcgt ccaccagcg cgccgtcctg 1560
gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgacccttg ccggacctca tccacccag gacgggcccgc 1680
ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
gcccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac 1920
acggagaccg ccagctggat gttcggcgtc cccggggagg ccgtggaccc cctgatgcgc 1980
cgggcggcca agaccatcaa cttcggggtc ctctacggca tgcggccca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
gtggagaccc tcttcggccg ccgcgcctac gtgccagacc tagaggcccg ggtgaagagc 2220
gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
ctcatgaagc tggctatggt gaagctcttc ccaggctgg aggaaatggg ggccaggatg 2340
ctccttcagg tccacaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
gcccggctgg ccaaggagggt catggagggg gtgtatcccc tggccgtgcc cctggagggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggag accaccacca ccaccac 2517

<210> 77
<211> 2526
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 77
atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60

| | | | | | | |
|------------|------------|-------------|-------------|------------|-------------|------|
| ggccaccacc | tggcctaccg | cacctttcttc | gccctgaagg | gcctcaccac | gagccggggc | 120 |
| gaaccggtgc | aggcggtcta | cggcttcgcc | aagagcctcc | tcaaggccct | gaaggaggac | 180 |
| gggtacaagg | ccgtcttcgt | ggtctttgac | gccaaaggccc | cctccttcgg | ccacgaggcc | 240 |
| tacgaggcct | acaaggcggg | gagggccccc | accccgagg | acttcccccg | gcagctcgcc | 300 |
| ctcatcaagg | agctggtgga | cctcctgggg | tttaccgcgc | tcgaggtccc | cggctacgag | 360 |
| gcggacgacg | ttctcgccac | cctggccaag | aaggcggaaa | aggaggggta | cgaggtgcgc | 420 |
| atcctcaccg | ccgaccgcga | cctctaccaa | ctcgtctccg | accgcgtcgc | cgtcctccac | 480 |
| cccgagggcc | acctcatcac | cccggagtgg | ctttgggaga | agtacggcct | caggccggag | 540 |
| cagtgggtgg | acttccgcgc | cctcgtgggg | gaccctccg | acaacctccc | cggggtcaag | 600 |
| ggcatcgggg | agaagaccgc | cctcaagctc | ctcaaggagt | ggggaagcct | ggaaaacctc | 660 |
| ctcaagaacc | tggaccgggt | aaagccagaa | aacgtccggg | agaagatcaa | ggccacctg | 720 |
| gaagacctca | ggctctcctt | ggagctctcc | cgggtgcgca | ccgacctccc | cctggagggtg | 780 |
| gacctcgccc | aggggcggga | gcccgaccgg | gaggggctta | gggccttcct | ggagagggtg | 840 |
| gagttcgcca | gcctcctcca | cgagttcggc | ctcctggagg | ccccgcccc | cctggaggag | 900 |
| gccccctggc | ccccgcggga | aggggccttc | gtgggcttcg | tcctctcccg | ccccgagccc | 960 |
| atgtgggcgg | agcttaaagc | cctggccgcc | tgcagggggc | gccgcgtgca | ccgggcagca | 1020 |
| gacccttggt | cggggctaaa | ggacctcaag | gaggtccggg | gcctcctcgc | caaggacctc | 1080 |
| gccgtcttgg | cctcgaggga | ggggctagac | ctcgtgcccc | gggacgacct | catgctcctc | 1140 |
| gcctacctcc | tggacccttc | gaacaccacc | cccgaggggg | tggcgcgggc | ctacgggggg | 1200 |
| gagtggacgg | aggacgccgc | ccaccggggc | ctcctctcgg | agaggctcca | tcggaacctc | 1260 |
| cttaagcgcc | tcgaggggga | ggagaagctc | ctttggctct | accacgaggt | ggaaaagccc | 1320 |
| ctctcccggg | tcctggccca | tatggaggcc | accggggtag | ggcgggacgt | ggcctacctt | 1380 |
| caggcccttt | ccctggagct | tgcggaggag | atccgcgcgc | tcgaggagga | ggtcttccgc | 1440 |
| ttggcggggc | accccttcaa | cctcaactcc | cgggaccagc | tggaaagggt | gctctttgac | 1500 |
| gagcttaggc | ttcccgccct | ggggaagacg | caaaagacag | gcaagcgctc | caccagcgcc | 1560 |
| gcggtgctgg | aggccctacg | ggaggcccac | cccatcgtgg | agaagatcct | ccagcaccgg | 1620 |
| gagctcacca | agctcaagaa | cacctacgtg | gacccctcc | caagcctcgt | ccacccgagg | 1680 |
| acgggcgcgc | tccacacccg | cttcaaccag | acggccacgg | ccacggggag | gcttagtagc | 1740 |

| | |
|--|------|
| tccgacccca acctgcagaa catccccgtc cgcaccccct tgggccagag gatccgccgg | 1800 |
| gccttcatcg ccgaggaggg gtggctattg gtggccctgg actatagcca gatagagctc | 1860 |
| agggtgctgg cccacctctc cggcgacgag aacctgatcc gggctctcca ggaggggagg | 1920 |
| gacatccaca cggagaccgc cagctggatg ttcggcgctcc cccgggaggc cgtggacccc | 1980 |
| ctgatgcgcc gggcggccaa gaccatcaac ttcgggggtcc tctacggcat gtcggccccc | 2040 |
| cgcctctccc aggagctagc catcccttac gaggaggccc aggccttcat tgagcgctac | 2100 |
| tttcagagct tccccagggt gcgggcctgg attgagaaga ccctggagga gggcaggagg | 2160 |
| cgggggtacg tggagaccct ctctggccgc cgccgctacg tgccagacct agaggcccgg | 2220 |
| gtgaagagcg tgcgggaggc ggccgagcgc atggccttca acatgcccgt ccagggcacc | 2280 |
| gccgccgacc tcatgaagct ggctatggtg aagctcttcc ccaggctgga ggaaatgggg | 2340 |
| gccaggatgc tccttcagggt ccacaacgag ctggctctcg agggcccaaa agagagggcg | 2400 |
| gaggccgtgg cccggctggc caaggaggtc atggaggggg tgtatcccct ggccgtgccc | 2460 |
| ctggaggtgg aggtggggat aggggaggac tggctctccg ccaaggagca ccaccaccac | 2520 |
| caccac | 2526 |

<210> 78
 <211> 2526
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

| | |
|--|-----|
| <400> 78 | |
| atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac | 60 |
| ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc | 120 |
| gaaccggtgc aggcggtcta cggtctcgcc aagagcctcc tcaaggccct gaaggaggac | 180 |
| gggtacaagg ccgtcttcgt ggtctttgac gccaaaggccc cctccttccg ccacgaggcc | 240 |
| tacgaggcct acaaggcggg gagggccccg acccccagagg acttcccccg gcagctcgcc | 300 |
| ctcatcaagg agctggtgga cctcctgggg ttaccgcgc tcgaggtccc cggctacgag | 360 |
| gcggacgacg ttctcgccac cctggccaag aaggcggaaa aggaggggta cgaggtgcgc | 420 |
| atcctcaccg ccgaccgcga cctctaccaa ctcgctctccg accgcgtcgc cgtcctccac | 480 |
| cccagaggcc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag | 540 |
| cagtgggtgg acttccgcgc cctcgtgggg gaccctccg acaacctccc cgggggtcaag | 600 |

| | | | | | | |
|------------|-------------|-------------|------------|------------|-------------|------|
| ggcatcgggg | agaagaccgc | cctcaagctc | ctcaaggagt | ggggaagcct | ggaaaacctc | 660 |
| ctcaagaacc | tggaccgggt | aaagccagaa | aacgtccggg | agaagatcaa | ggcccacctg | 720 |
| gaagacctca | ggctctcctt | ggagctctcc | cgggtgcgca | ccgacctccc | cctggagggtg | 780 |
| gacctcgccc | aggggcggga | gcccgaccgg | gaggggctta | gggccttcct | ggagaggctg | 840 |
| gagttcggca | gcctcctcca | cgagttcggc | ctcctggagg | ccccgcgcc | cctggaggag | 900 |
| gccccctggc | ccccgcggga | aggggccttc | gtgggcttcg | tcctctcccg | ccccgagccc | 960 |
| atgtggggcg | agcttaaagc | cctggccggc | tgcagggggc | gccgcgtcca | ccggggcccc | 1020 |
| gagccttata | aagccctcag | ggacctgaag | gaggcgcggg | ggcttctcgc | caaagacctg | 1080 |
| agcgttcttg | ccctgaggga | aggccttggc | ctccgcgccg | gcgacgacct | catgctcctc | 1140 |
| gcctacctcc | tggacccttc | gaacaccacc | cccagggggg | tggcccggcg | ctacggcggg | 1200 |
| gagtggacgg | aggaggcggg | ggagcgggcc | gccctttccg | agaggctctt | cgccaacctg | 1260 |
| tgggggaggg | ttgaggggga | ggagaggctc | ctttggcttt | accgggaggt | ggagaggccc | 1320 |
| ctttccgctg | tcctggccca | tatggaggcc | acgggggtgc | gcctggacgt | ggcctatctc | 1380 |
| agggccttgt | ccctggagggt | ggccgaggag | atcgcccgcc | tcgaggccga | ggtcttccgc | 1440 |
| ctggccggcc | accccttcaa | cctcaactcc | cgggaccagc | tggaaagggt | cctctttgac | 1500 |
| gagctagggc | ttcccgccat | cggcaagacg | gagaagaccg | gcaagcgctc | caccagcgcc | 1560 |
| gccgtcctgg | aggccctccg | cgaggcccac | cccatcggtg | agaagatcct | gcagtaccgg | 1620 |
| gagctcacca | agctgaagag | cacctacatt | gacctcttgc | cggacctcat | ccaccccagg | 1680 |
| acgggccgcc | tccacacccg | cttcaaccag | acggccacgg | ccacgggcag | gctaagtagc | 1740 |
| tccgatccca | acctccagaa | catccccgtc | cgcaccccg | ttgggcagag | gatccgcggg | 1800 |
| gccttcgtgg | ccgaggcggg | ttgggcgttg | gtggccctgg | actatagcca | gatagagctc | 1860 |
| cgcgtcctcg | cccacctctc | cggggacgaa | aacctgatca | gggtcttcca | ggaggggaag | 1920 |
| gacatccaca | cccagaccgc | aagctggatg | ttcggcgctc | ccccggaggc | cgtggacccc | 1980 |
| ctgatgcgcc | gggcggccaa | gacgggtgaac | ttcggcgctc | tctacggcat | gtccgcccac | 2040 |
| aggctctccc | aggagcttgc | catcccctac | gaggaggcgg | tggcctttat | agagcgctac | 2100 |
| ttccaaagct | tccccaaagg | gcgggcctgg | atagaaaaga | ccctggagga | ggggaggaag | 2160 |
| cggggctacg | tggaaaccct | cttcggaaga | aggcgctacg | tgcccgacct | caacgcccgg | 2220 |
| gtgaagagcg | tcagggaggc | cgcggagcgc | atggccttca | acatgcccgt | ccagggcacc | 2280 |

| | | | | | | |
|------------|------------|------------|------------|------------|------------|------|
| gccgccgacc | tcataaagct | cgccatggtg | aagctcttcc | cccgcctccg | ggagatgggg | 2340 |
| gcccgcacgc | tcctccaggt | ccacaacgag | ctcctcctgg | aggcccccca | agcgcgggcc | 2400 |
| gaggaggtgg | cggttttggc | caaggaggcc | atggagaagg | cctatcccct | cgccgtgccc | 2460 |
| ctggaggtgg | aggtggggat | gggggaggac | tggttttccg | ccaagggtca | ccaccaccac | 2520 |
| caccac | | | | | | 2526 |

<210> 79
 <211> 2499
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

| | |
|------------|---|
| <400> 79 | |
| atgaattcgg | ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60 |
| caccacctgg | cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120 |
| ccggtgcagg | cggtctacgg cttcgccaag agcctcctca aggcctcaa ggaggacggg 180 |
| gacgcggtga | tcgtggtcct tgacgccaaag gccccctcct tccgccacga ggcctacggg 240 |
| gggtacaagg | cgggccgggg ccccaacgcc gaggactttc ccgggcaact cgccctcatc 300 |
| aaggagctgg | tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360 |
| gacgtcctgg | ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcacctc 420 |
| accgccgaca | aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccag 480 |
| gggtacctca | tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540 |
| gccgactacc | gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc 600 |
| ggggagaaga | cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660 |
| aacctggacc | ggctgaagcc cgccatccgg gagaagatcc tggccacat ggacgatctg 720 |
| aagctctcct | gggacctggc caagggtgcg accgacctgc ccctggaggt ggacttcgcc 780 |
| aaaaggcggg | agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840 |
| agcctcctcc | acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900 |
| cccccgccgg | aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtggggc 960 |
| gatcttctgg | ccctggccgc cgccaggggc ggccgcgtgc accgggcagc agacccttg 1020 |
| gcggggctaa | aggacctcaa ggaggtccgg ggcctcctcg ccaaggacct cgccgtcttg 1080 |
| gcctcgaggg | aggggctaga cctcgtgccc ggggacgacc ccatgctcct cgcctacctc 1140 |

ctggaccctt cgaacaccac ccccgagggg gtggcgcggc gctacggggg ggagtggacg 1200
gaggacgccg cccaccgggc cctcctctcg gagaggctcc atcggaacct ccttaagcgc 1260
ctcgaggggg aggagaagct cctttggctc taccacgagg tggaaaagcc cctctcccgg 1320
gtcctggccc atatggaggc caccggggta cggcgggacg tggcctacct tcaggccctt 1380
tccctggagc ttgcggagga gatccgccgc ctcgaggagg aggtcttccg cttggcgggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tgctctttga cgagcttagg 1500
cttcccgcct tggggaagac gcaaaagaca ggcaagcgt ccaccagcgc cgcggtgctg 1560
gaggccctac gggaggccca ccccatcgtg gagaagatcc tccagcaccg ggagctcacc 1620
aagctcaaga acacctacgt ggacccccctc ccaagcctcg tccacccgag gacgggcccgc 1680
ctccacaccc gcttcaacca gacggccacg gccacgggga ggcttagtag ctccgacccc 1740
aacctgcaga acatccccgt ccgcaccccc ttgggcccaga ggatccgccg ggccttcgtg 1800
gccgaggcgg gttgggcgtt ggtggccctg gactatagcc agatagagct ccgcgtcctc 1860
gcccacctct ccggggacga aaacctgatc agggctcttc aggaggggaa ggacatccac 1920
acccagaccg caagctggat gttcggcgtc ccccgagg cctgggacct cctgatgcgc 1980
cgggcggcca agacggtgaa cttcggcgtc ctctacggca tgtccgcca taggctctcc 2040
caggagcttg ccatccccta cgaggaggcg gtggccttta tagagcgcta cttccaaagc 2100
ttccccaagg tgcgggcctg gatagaaaag accctggagg aggggaggaa gcggggctac 2160
gtggaaaacc tcttcggaag aaggcgctac gtgcccgaac tcaacgcccg ggtgaagagc 2220
gtcagggagg ccgcggagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
ctcatgaagc tcgccatggt gaagctcttc cccgcctcc gggagatggg ggcccgcacg 2340
ctcctccagg tccacgacga gctcctcctg gaggccccc aagcgcgggc cgaggaggtg 2400
gcggctttgg ccaaggaggc catggagaag gcctatcccc tcgccgtgcc cctggaggtg 2460
gaggtgggga tgggggagga ctggctttcc gccaaagggt 2499

<210> 80
<211> 2499
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 80

| | | | | | | |
|-------------|------------|-------------|------------|------------|------------|------|
| atgaattcgg | ggatgctgcc | cctctttgag | cccaagggcc | gggtcctcct | ggtggacggc | 60 |
| caccacctgg | cctaccgcac | cttccacgcc | ctgaagggcc | tcaccaccag | ccggggggag | 120 |
| ccggtgcagg | cggtctacgg | cttgcgcaag | agcctcctca | aggccctcaa | ggaggacggg | 180 |
| gacgcggtga | tcgtggtctt | tgacgccaag | gccccctcct | tccgccacga | ggcctacggg | 240 |
| gggtacaagg | cgggccgggg | ccccacgccg | gaggactttc | cccggcaact | cgccctcatc | 300 |
| aaggagctgg | tggacctcct | ggggctggcg | cgctctgagg | tcccgggcta | cgaggcggac | 360 |
| gacgtcctgg | ccagcctggc | caagaaggcg | gaaaaggagg | gctacgaggt | ccgcctcctc | 420 |
| accgccgaca | aagaccttta | ccagctcctt | tccgaccgca | tccacgtcct | ccaccccag | 480 |
| gggtacctca | tcaccccggc | ctggcttttg | gaaaagtacg | gcctgaggcc | cgaccagtgg | 540 |
| gccgactacc | gggccctgac | cggggacgag | tccgacaacc | ttcccggggg | caagggcatc | 600 |
| ggggagaaga | cggcgaggaa | gcttctggag | gagtggggga | gcctggaagc | cctcctcaag | 660 |
| aacctggacc | ggctgaagcc | cgccatccgg | gagaagatcc | tggcccacat | ggacgatctg | 720 |
| aagctctcct | gggacctggc | caaggtgcgc | accgacctgc | ccctggaggt | ggacttcgcc | 780 |
| aaaaggcggg | agcccgaccg | ggagaggctt | agggcctttc | tggagaggct | tgagtttggc | 840 |
| agcctcctcc | acgagttcgg | ccttctggaa | agccccaagg | ccctggagga | ggccccctgg | 900 |
| ccccgcgagg | aaggggcctt | cgtgggcttt | gtgctttccc | gcaaggagcc | catgtggggc | 960 |
| gatcttctgg | ccctggccgc | cgccaggggc | ggccgcgtgc | accgggcagc | agacccttg | 1020 |
| gcgggggctaa | aggacctcaa | ggaggtccgg | ggcctcctcg | ccaaggacct | cgccgtcttg | 1080 |
| gcctcgaggg | aggggctaga | cctcgtgccc | ggggacgacc | ccatgctcct | cgctacctc | 1140 |
| ctggaccctt | cgaacaccac | ccccgagggg | gtggcgcggc | gctacggggg | ggagtggacg | 1200 |
| gaggacgccg | cccaccgggc | cctcctctcg | gagaggctcc | atcggaacct | ccttaagcgc | 1260 |
| ctcgaggggg | aggagaagct | cctttggctc | taccacgagg | tggaaaagcc | cctctcccgg | 1320 |
| gtcctggccc | atatggaggc | caccggggta | cggcgggacg | tggcctacct | tcaggccctt | 1380 |
| tccctggagc | ttgcggagga | gatccgccgc | ctcgaggagg | aggtcttccg | cttggcgggc | 1440 |
| cacccttca | acctcaactc | ccgggaccag | ctggaaaggg | tgctctttga | cgagcttagg | 1500 |
| cttcccgcct | tggggaagac | gcaaaagaca | ggcaagcgct | ccaccagcgc | cgcggtgctg | 1560 |
| gaggccctac | gggaggccca | ccccatcgct | gagaagatcc | tccagcaccg | ggagctcacc | 1620 |
| aagctcaaga | acacctacgt | ggacccccctc | ccaagcctcg | tccacccgag | gacgggccgc | 1680 |
| ctccacaccc | gcttcaacca | gacggccacg | gccacgggga | ggcttagtag | ctccgacccc | 1740 |

| | |
|--|------|
| aacctgcaga acatccccgt ccgcaccccc ttggggccaga ggatccgccg ggccttcac | 1800 |
| gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct caggggtgctg | 1860 |
| gcccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggagc ggacatccac | 1920 |
| acggagaccg ccagctggat gttcggcgtc ccccgaggag ccgtggaccc cctgatgcgc | 1980 |
| cgggcggcca agaccatcaa cttcggggtc ctctacggca tgcgggcca ccgcctctcc | 2040 |
| caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc | 2100 |
| ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac | 2160 |
| gtggagaccc tcttcggccg ccgcccgtac gtgccagacc tagaggcccg ggtgaagagc | 2220 |
| gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac | 2280 |
| ctcatgaagc tggctatggt gaagctcttc cccaggctgg aggaaatggg ggccaggatg | 2340 |
| ctccttcagg tccacgacga gctggctctc gaggcccaa aagagagggc ggaggccgtg | 2400 |
| gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg | 2460 |
| gaggtgggga taggggagga ctggctctcc gccaaaggag | 2499 |

<210> 81
 <211> 2499
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

| | |
|---|-----|
| <400> 81 | |
| atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggctctcct ggtggacggc | 60 |
| caccacctgg cctaccgcac ctccaacgcc ctgaagggcc tcaccaccag ccggggggag | 120 |
| ccggtgcagg cggctacagg ctccgccaag agcctctca aggcctcaa ggaggacggg | 180 |
| gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg | 240 |
| gggtacaagg cgggccgggc cccacgccg gaggacttcc cccggcaact cgccctcatc | 300 |
| aaggagctgg tggacctcct ggggctggcg cgctcgagg tcccgggcta cgaggcggac | 360 |
| gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctctc | 420 |
| accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag | 480 |
| gggtacctca tcaccccggc ctggcttttg gaaaagtaag gcctgaggcc cgaccagtgg | 540 |
| gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggt caagggcac | 600 |

| | | | | | | |
|------------|------------|------------|------------|------------|------------|------|
| ggggagaaga | cggcgaggaa | gcttctggag | gagtggggga | gcctggaagc | cctcctcaag | 660 |
| aacctggacc | ggctgaagcc | cgccatccgg | gagaagatcc | tggcccacat | ggacgatctg | 720 |
| aagctctcct | gggacctggc | caaggtgcg | accgacctgc | ccctggaggt | ggacttcgcc | 780 |
| aaaaggcggg | agcccgaccg | ggagaggctt | agggcctttc | tggagaggct | tgagtttggc | 840 |
| agcctcctcc | acgagttcgg | ccttctggaa | agccccaagg | ccctggagga | ggccccctgg | 900 |
| ccccgcgcg | aaggggcctt | cgtgggcttt | gtgctttccc | gcaaggagcc | catgtgggcc | 960 |
| gatcttctgg | ccctggccgc | cgccaggggc | ggccgcgtcc | accgggcccc | cgagccttat | 1020 |
| aaagccctca | gggacctgaa | ggaggcgcg | gggcttctcg | ccaaagacct | gagcgttctg | 1080 |
| gccctgaggg | aaggccttgg | cctcccgccc | ggcgacgacc | ccatgctcct | cgctacctc | 1140 |
| ctggaccctt | cgaacaccac | ccccgagggg | gtggccccgc | gctacggcgg | ggagtggacg | 1200 |
| gaggaggcgg | gggagcgggc | cgccctttcc | gagaggctct | tcgccaacct | gtgggggagg | 1260 |
| cttgaggggg | aggagaggct | cctttggctt | taccgggagg | tggagaggcc | cctttccgct | 1320 |
| gtcctggccc | atatggaggc | cacgggggtg | cgcttgacg | tggcctatct | cagggccttg | 1380 |
| tccttgagg | tggccgagga | gatcgccgc | ctcgaggccg | aggtcttccg | cctggccggc | 1440 |
| cacccttca | acctcaactc | ccgggaccag | ctggaaaagg | tcctctttga | cgagctaggg | 1500 |
| cttcccgcca | tcggcaagac | ggagaagacc | ggcaagcgct | ccaccagcgc | cgccgtcctg | 1560 |
| gaggccctcc | gcgaggccca | ccccatcgtg | gagaagatcc | tgcagtaccg | ggagctcacc | 1620 |
| aagctgaaga | gcacctacat | tgacccttg | ccggacctca | tccaccccag | gacgggccgc | 1680 |
| ctccacaccc | gcttcaacca | gacggccacg | gccacgggca | ggctaagtag | ctccgatccc | 1740 |
| aacctccaga | acatccccgt | ccgcaccccg | cttgggcaga | ggatccgccg | ggccttcgtg | 1800 |
| gccgaggcgg | gttgggcgtt | ggtggccctg | gactatagcc | agatagagct | ccgcgtcctc | 1860 |
| gcccacctct | ccggggacga | aaacctgac | agggctcttc | aggaggggaa | ggacatccac | 1920 |
| accagaccg | caagctggat | gttcggcgtc | ccccggagg | ccgtggaccc | cctgatgcgc | 1980 |
| cgggcggcc | agacggtgaa | cttcggcgtc | ctctacggca | tgtccgcca | taggctctcc | 2040 |
| caggagcttg | ccatccccta | cgaggaggcg | gtggccttta | tagagcgcta | cttccaaagc | 2100 |
| ttccccaagg | tgcgggcctg | gatagaaaag | accctggagg | aggggaggaa | gcggggctac | 2160 |
| gtggaaaccc | tcttcggaag | aaggcgctac | gtgcccgacc | tcaacgcccg | ggtgaagagc | 2220 |
| gtcagggagg | ccgcggagcg | catggccttc | aacatgcccg | tccagggcac | cgccgccgac | 2280 |
| ctcatgaagc | tcgccatggt | gaagctcttc | ccccgcctcc | gggagatggg | ggccccgatg | 2340 |

ctcctccagg tccacgacga gctcctcctg gaggccccc aagcgcgggc cgaggaggtg 2400
 gcggcttttg ccaaggaggc catggagaag gcctatcccc tcgccgtgcc cctggaggtg 2460
 gaggtgggga tgggggagga ctggctttcc gccaaaggt 2499

<210> 82
 <211> 2499
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 82
 atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
 caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
 ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
 gacgcggtga tcgtggtctt tgacgccaag gccccctcct tcgcccacga ggcctacggg 240
 ggggtacaagg cgggcccggc cccacgccc gaggactttc cccggcaact cgcctcatc 300
 aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
 gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatcctc 420
 accgccgaca aagaccttta ccagctcctt tcgaccgca tcacgtcct ccaccccag 480
 gggtagctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
 gccgactacc gggccctgac cggggacgag tcgacaacc ttcccggggc caagggcatc 600
 ggggagaaga cggcgaggaa gcttctggag gagggggga gcctggaagc cctcctcaag 660
 aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccatat ggacgatctg 720
 aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780
 aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
 agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
 cccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc 960
 gatcttcttg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
 aaagccctca gggacctgaa ggaggcgcg gggcttctcg ccaaagacct gagcgttctg 1080
 gccctgaggg aaggccttgg cctcccgcc ggcgacgacc ccatgctcct cgcctacctc 1140
 ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200

| | | | | | | |
|------------|------------|------------|------------|------------|-------------|------|
| gaggaggcgg | gggagcgggc | cgccctttcc | gagaggctct | tcgccaacct | gcttaagagg | 1260 |
| cttgaggggg | aggagaggct | cctttggctt | taccgggagg | tggagaggcc | cctttccgct | 1320 |
| gtcctggccc | atatggaggc | cacgggggtg | cgcttggaac | tggcctatct | cagggccttg | 1380 |
| tccctggagg | tggccgagga | gatcgcccg | ctcgaggccg | aggtcttccg | cctggccggc | 1440 |
| cacccttca | acctcaactc | cggggaccag | ctggaaaggg | tcctctttga | cgagctaggg | 1500 |
| cttcccgcca | tcggaagac | gcaaaagacc | ggcaagcgct | ccaccagcgc | cgccgtcctg | 1560 |
| gaggccctcc | gcgaggccca | ccccatcgtg | gagaagatcc | tgcagtaccg | ggagctcacc | 1620 |
| aagctgaaga | gcacctacat | tgacctcttg | ccggacctca | tccaccccag | gacgggccgc | 1680 |
| ctccacaccc | gcttcaacca | gacggccacg | gccacgggca | ggctaagtag | ctccgatccc | 1740 |
| aacctccaga | acatccccgt | ccgcaccccg | cttgggcaga | ggatccgccg | ggccttcctc | 1800 |
| gccgaggagg | ggtggctatt | ggtggccctg | gactatagcc | agatagagct | cagggtgctg | 1860 |
| gcccacctct | ccggcgacga | gaacctgatc | cggtctcttc | aggagggggc | ggacatccac | 1920 |
| acggagaccg | ccagctggat | gttcggcgtc | ccccgggagg | ccgtggaccc | cctgatgcgc | 1980 |
| cgggcggcc | agaccatcaa | cttcggggtc | ctctacggca | tgtcggccca | ccgcctctcc | 2040 |
| caggagctag | ccatccctta | cgaggaggcc | caggccttca | ttgagcgcta | ctttcagagc | 2100 |
| ttccccaagg | tgcgggcctg | gattgagaag | accctggagg | agggcaggag | gcgggggtac | 2160 |
| gtggagaccc | tcttcggccg | ccgccgctac | gtgccagacc | tagaggcccg | ggtgaagagc | 2220 |
| gtgcgggagg | cggccgagcg | catggccttc | aacatgcccg | tccagggcac | cgccgccgac | 2280 |
| ctcatgaagc | tggctatggt | gaagctcttc | cccaggettg | aggaaatggg | ggccaggatg | 2340 |
| ctccttcagg | tccacgacga | gctggtcctc | gaggccccaa | aagagagggc | ggaggccgtg | 2400 |
| gcccggtgg | ccaaggaggt | catggagggg | gtgtatcccc | tggccgtgcc | cctggagggtg | 2460 |
| gaggtgggga | taggggagga | ctggctctcc | gccaaaggag | | | 2499 |

| | |
|-------|------------|
| <210> | 83 |
| <211> | 2517 |
| <212> | DNA |
| <213> | Artificial |

<220>
<223> Synthetic

<400> 83
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120

| | |
|--|------|
| ccggtgcagg cggtctacgg ctctgccaag agcctcctca aggccctcaa ggaggacggg | 180 |
| gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccaaga ggccctacggg | 240 |
| gggtacaagg cgggccgggc cccacgcgcg gaggactttc cccggcaact cgcctcatc | 300 |
| aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac | 360 |
| gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc | 420 |
| accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag | 480 |
| gggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg | 540 |
| gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcac | 600 |
| ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag | 660 |
| aacctggacc ggctgaagcc cgccatcccg gagaagatcc tggcccatat ggacgatctg | 720 |
| aagctctcct gggacctggc caagggtgcg accgacctgc ccctggagggt ggacttcgcc | 780 |
| aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtgtggc | 840 |
| agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg | 900 |
| cccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc | 960 |
| gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat | 1020 |
| aaagccctca gggacctgaa ggaggcgcg gggtcttctc ccaaagacct gagcgttctg | 1080 |
| gccctgaggg aaggccttgg cctcccgcgc ggcgacgacc ccatgctcct cgcctacctc | 1140 |
| ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcg ggagtggacg | 1200 |
| gaggaggcgg ggcaccgggc cgccctttcc gagaggctct tcgccaacct gtgggggagg | 1260 |
| cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct | 1320 |
| gtcctggccc atatggaggc caccggggta cggcgggacg tggcctacct tcaggccctt | 1380 |
| tccctggagc ttgcggagga gatccgcgc ctcgaggagg aggtcttccg cttggcgggc | 1440 |
| cacccttca acctcaactc ccgggaccag ctggaaaggg tgctctttga cgagcttagg | 1500 |
| cttccgcgct tggggaagac gcaaaagaca ggcaagcgct ccaccagcgc cgcggtgctg | 1560 |
| gaggccctac gggaggccca cccatcgctg gagaagatcc tccagcaccg ggagctcacc | 1620 |
| aagctcaaga acacctacgt ggacccccctc ccaagcctcg tccacccgag gacgggcgc | 1680 |
| ctccacaccc gcttcaacca gacggccacg gccacgggga ggcttagtag ctccgacccc | 1740 |
| aacctgcaga acatccccgt ccgcaccccc ttgggcccaga ggatccgcgc ggccttcac | 1800 |

gcccaggagg ggtggctatt ggtggccctg gactatagcc agatagagct caggggtgctg 1860
 gccacactct ccggcgacga gaacctgatc cgggtcttcc aggaggggag ggacatccac 1920
 acggagaccg ccagctggat gttcggcgtc ccccgaggag ccgtggaccc cctgatgcgc 1980
 cgggcggcca agaccatcaa cttcggggtc ctctacggca tgcggccca ccgcctctcc 2040
 caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
 gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
 ctcatgaagc tggctatggt gaagctcttc cccaggctgg aggaaatggg ggccaggatg 2340
 ctcttcagg tccacaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
 gcccggttg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggagggtg 2460
 gaggtgggga taggggagga ctggctctcc gccaggagc accaccacca ccaccac 2517

<210> 84
 <211> 2517
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 84
 atgaattcgg ggatgctgcc cctctttgag cccaagggcc ggtcctcct ggtggacggc 60
 caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
 ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggcctcaa ggaggacggg 180
 gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
 ggggtacaagg cgggccgggc ccccaacgag gaggacttct cccggcaact cgccctcatc 300
 aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
 gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420
 accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
 gggtagctca tcaccccgcc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
 gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggt caagggcac 600
 ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
 aacctggacc ggctgaagcc cgccatcccg gagaagatcc tggcccatat ggacgatctg 720

| | | |
|-----------------------------------|-----------------------------------|------|
| aaagctctcct gggacctggc caaggtgcg | accgacctgc ccctggaggt ggacttcgcc | 780 |
| aaaaggcggg agcccgaccg ggagaggctt | agggcctttc tggagaggct tgagtttggc | 840 |
| agcctcctcc acgagttcgg ccttctggaa | agccccaagg ccctggagga ggccccctgg | 900 |
| cccccgccgg aaggggcctt cgtgggcttt | gtgctttccc gcaaggagcc catgtgggcc | 960 |
| gatcttctgg ccctggccgc cgccaggggc | ggccgcgtcc accgggcccc cgagccttat | 1020 |
| aaagccctca gggacctgaa ggaggcgcg | gggcttctcg ccaaagacct gagcgttctg | 1080 |
| gccctgaggg aaggccttgg cctcccgccc | ggcgacgacc ccatgctcct cgccctaccc | 1140 |
| ctggaccctt cgaacaccac ccccgagggg | gtggcccggc gctacggcgg ggagtggacg | 1200 |
| gaggaggcgg gggagcgggc cgccctttcc | gagaggctcc atcggaacct gtgggggagg | 1260 |
| cttgaggggg aggagaggct cctttggctt | taccgggagg tggagaggcc cctttccgct | 1320 |
| gtcctggccc atatggaggc caccggggta | cggcgggacg tggcctacct tcaggccctt | 1380 |
| tccctggagc ttgcggagga gatccgccgc | ctcgaggagg aggtcttccg cttggcgggc | 1440 |
| caccccttca acctcaactc ccgggaccag | ctggaaaggg tgctctttga cgagcttagg | 1500 |
| cttcccgctt tggggaagac gcaaaagaca | ggcaagcgtt ccaccagcgc cgcggtgctg | 1560 |
| gaggccctac gggaggccca ccccatcgctg | gagaagatcc tccagcaccg ggagctcacc | 1620 |
| aagctcaaga acacctacgt ggacccccctc | ccaagcctcg tccacccgag gacgggcccgc | 1680 |
| ctccacaccc gcttcaacca gacggccacg | gccacgggga ggcttagtag ctccgacccc | 1740 |
| aacctgcaga acatccccgt ccgcaccccc | ttggggcaga ggatccgccg ggccttcctc | 1800 |
| gccgaggagg ggtggctatt ggtggccctg | gactatagcc agatagagct caggggtgctg | 1860 |
| gcccacctct ccggcgacga gaacctgatc | cgggctcttc aggaggggcg ggacatccac | 1920 |
| acggagaccg ccagctggat gttcggcgtc | ccccgggagg ccgtggaccc cctgatgcgc | 1980 |
| cgggcgggca agaccatcaa cttcggggtc | ctctacggca tgtcggccca ccgcctctcc | 2040 |
| caggagctag ccatccctta cgaggaggcc | caggccttca ttgagcgcta ctttcagagc | 2100 |
| ttccccaagg tgcgggcctg gattgagaag | accctggagg agggcaggag gcgggggtac | 2160 |
| gtggagaccc tcttcggccg ccgcccgtac | gtgccagacc tagaggcccg ggtgaagagc | 2220 |
| gtgcgggagg cgcccgagcg catggccttc | aacatgcccc tccagggcac cgccgccgac | 2280 |
| ctcatgaagc tggctatggt gaagctcttc | cccaggctgg aggaaatggg ggccaggatg | 2340 |
| ctccttcagg tccacaacga gctggtcctc | gaggccccaa aagagagggc ggaggccgtg | 2400 |

gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 85
<211> 2517
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 85
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
ccgggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcgggtga tcgtgggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cggggccgggc cccacgccc gaggaacttc cccggcaact cgcctcatc 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcacctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccggc ctggctttgg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggc caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg 720
aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
ccccgcggg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtggggc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
aaagccctca gggacctgaa ggaggcggcg gggcttctcg ccaaagacct gagcgttctg 1080
gccctgaggg aaggccttgg cctcccggcc ggcgacgacc ccatgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200
gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttaagagg 1260
cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct 1320

gtcctggccc atatggaggc caccggggta cggcgggacg tggcctacct tcaggccctt 1380
tccctggagc ttgcggagga gatccgccgc ctcgaggagg aggtcttccg cttggcgggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tgctctttga cgagcttagg 1500
cttccgcct tggggaagac gcaaaagaca ggcaagcgct ccaccagcgc cgcggtgctg 1560
gaggccctac gggaggccca ccccatcgtg gagaagatcc tccagcaccg ggagctcacc 1620
aagctcaaga acacctacgt ggacccctc ccaagcctcg tccaccgag gacgggcccgc 1680
ctccacacc gttcaacca gacggccacg gccacgggga ggcttagtag ctccgacccc 1740
aacctgcaga acatccccgt ccgaccccc ttgggccaga ggatccgccc ggcccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
gcccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac 1920
acggagaccg ccagctggat gttcggcgtc ccccgaggag ccgtggaccc cctgatgcgc 1980
cgggcggcca agaccatcaa cttcggggtc ctctacggca tgcggccca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccagg tgccggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
gtggagacc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
gtgcgggagg cggccgagcg catggccttc aacatgccc tccagggcac cgcgcgcgac 2280
ctcatgaagc tggctatggt gaagctcttc ccaggctgg aggaaatggg ggccaggatg 2340
ctccttcagg tccacaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
gagggtggga taggggagga ctggctctcc gccaggagc accaccacca ccaccac 2517

<210> 86
<211> 2517
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 86
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
ccggtgcagg cgggtctacg cttcgccaag agcctcctca aggcctcaa ggaggacggg 180

| | |
|---|------|
| gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg | 240 |
| gggtacaagg cgggcccgggc cccacgccg gaggactttc cccggcaact cgcctcatc | 300 |
| aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac | 360 |
| gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc | 420 |
| accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccag | 480 |
| gggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg | 540 |
| gccgactacc gggccctgac cggggacgag tccgacaacc ttcccgggggt caagggcatc | 600 |
| ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag | 660 |
| aacctggacc ggctgaagcc cgccatccg gagaaatcc tggcccacat ggacgatctg | 720 |
| aagctctcct gggacctggc caaggtgctc accgacctgc ccctggagggt ggacttcgcc | 780 |
| aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc | 840 |
| agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg | 900 |
| ccccgcggg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc | 960 |
| gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat | 1020 |
| aaagccctca gggacctgaa ggaggcgcg gggcttctcg ccaaagacct gagcgttctg | 1080 |
| gccctgaggg aaggccttgg cctcccgcgc ggcgacgacc ccatgctcct cgctacctc | 1140 |
| ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg | 1200 |
| gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gtgggggagg | 1260 |
| cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttcccgg | 1320 |
| gtcctggccc atatggaggc caccggggta cggcgggacg tggcctacct tcaggccctt | 1380 |
| tccttgagc ttgcggagga gatccgcgc ctcgaggagg aggtcttccg cttggcgggc | 1440 |
| caccccttca acctcaactc ccgggaccag ctggaaaggg tgctctttga cgagcttagg | 1500 |
| cttccgcct tggggaagac gcaaaagaca ggcaagcgct ccaccagcgc cgcggtgctg | 1560 |
| gaggccctac gggaggccca cccatcgtg gagaagatcc tccagcaccg ggagctcacc | 1620 |
| aagctcaaga acacctacgt ggacccccctc ccaagcctcg tccacccgag gacgggcccgc | 1680 |
| ctccacaccc gcttcaacca gacggccacg gccacgggga ggcttagtag ctccgacccc | 1740 |
| aacctgcaga acatccccgt ccgcaccccc ttggggcaga ggatccgcgc ggccttcac | 1800 |
| gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg | 1860 |
| gcccacctct ccggcgacga gaacctgatc cgggtcttcc aggagggggcg ggacatccac | 1920 |

| | |
|--|------|
| aaaaggcggg agccccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc | 840 |
| agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg | 900 |
| cccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc | 960 |
| gatcttctgg ccctggccgc cgccaggggc ggccgcgtgc accgggcagc agacccttg | 1020 |
| gcggggctaa aggacctcaa ggaggtccgg ggctcctcg ccaaggacct cgccgtcttg | 1080 |
| gcctcgaggg aggggctaga cctcgtgccc ggggacgacc ccatgctcct cgccctacctc | 1140 |
| ctggaccctt cgaacaccac ccccagggg gtggcgcggc gctacggggg ggagtggacg | 1200 |
| gaggacgccg cccaccgggc cctcctctcg gagaggctcc atcggaacct ccttaagcgc | 1260 |
| ctcgaggggg aggagaagct cctttggctc taccacgagg tggaaaagcc cctctcccgg | 1320 |
| gtcctggccc atatggaggc cacgggggtg cgccgggacg tggcctatct cagggccttg | 1380 |
| tccctggagg tggccgagga gatcgcccgc ctcgaggccg aggtcttccg cctggccggc | 1440 |
| cacccttca acctcaactc ccgggaccag ctggaaagg tctctttga cgagctaggg | 1500 |
| cttcccgcc tccggaagac ggagaagacc ggcaagcgt ccaccagcgc cgccgtcctg | 1560 |
| gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc | 1620 |
| aagctgaaga gcacctacat tgacccttg ccggacctca tccacccag gacgggccgc | 1680 |
| ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc | 1740 |
| aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggccttcac | 1800 |
| gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg | 1860 |
| gcccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac | 1920 |
| acggagaccg ccagctggat gttcggcgtc ccccgaggagg ccgtggacct cctgatgcgc | 1980 |
| cgggcggcca agaccatcaa cttcggggtc ctctacggca tgtcggcca ccgcctctcc | 2040 |
| caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc | 2100 |
| ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac | 2160 |
| gtggagacct tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc | 2220 |
| gtgcgggagg cgcccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac | 2280 |
| ctcatgaagc tggctatggt gaagctcttc ccaggtctgg aggaaatggg ggccaggatg | 2340 |
| ctccttcagg tccacaacga gctggtcctc gaggcccaaa aagagagggc ggaggccgtg | 2400 |
| gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggagggtg | 2460 |
| gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac | 2517 |

<210> 88
 <211> 2517
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 88
 atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
 caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag cgggggggag 120
 cgggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
 gacgcggtga tcgtggtcct tgacgccaag gccccctcct tccgccacga ggcctacggg 240
 ggggtacaagg cgggccgggc cccacgccc gaggactttc cccggcaact cgcctcctc 300
 aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
 gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420
 accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
 ggggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
 gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcctc 600
 ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
 aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccatat ggacgatctg 720
 aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggagggt ggacttcgcc 780
 aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
 agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
 ccccgccggg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc 960
 gatcttctgg ccctggccgc cgccaggggc ggccgcgtgc accgggcagc agacccttg 1020
 gcggggctaa aggacctcaa ggaggtccgg ggcctcctcg ccaaggacct cgccgtcttg 1080
 gcctcgaggg aggggctaga cctcgtgccc ggggacgacc ccatgctcct cgcctacctc 1140
 ctggaccctt cgaacaccac ccccgagggg gtggcgcggc gctacggggg ggagtggacg 1200
 gaggacgccg cccaccgggc cctcctctcg gagaggctcc atcggaacct ccttaagcgc 1260
 ctcgaggggg aggagaagct cctttggctc taccacgagg tggaaaagcc cctctcccgg 1320
 gtcttgcccc atatggaggc cacgggggtg cgctggacg tggcctatct ccaggccttg 1380

tccctggagg tggccgagga gatcgcccg ctcgaggccg aggtcttccg cctggccggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
cttcccgcga tcggcaagac ggagaagacc ggcaagcgct ccaccagcgc cgcggtcctg 1560
gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgaccccttg ccggacctca tccaccccag gacgggcccgc 1680
ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccc ggcccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
gccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac 1920
acggagaccg ccagctggat gttcggcgtc ccccgaggag ccgtggaccc cctgatgcgc 1980
cgggcggcca agaccatcaa cttcggggtc ctctacggca tgcggccca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgcgcgcgac 2280
ctcatgaagc tggctatggt gaagctcttc cccaggctgg aggaaatggg ggccaggatg 2340
ctccttcagg tccacaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
gcccggtgga ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggagggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 89
<211> 2517
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 89
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggctctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
ccggtgcagg cggctctacg ctctcgccaag agcctctctca aggcctctca ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggcccggc cccacgccc gaggacttct cccggcaact cgccctcatc 300

| | | | | | | |
|-------------|------------|------------|------------|------------|------------|------|
| aaggagctgg | tggacctcct | ggggctggcg | cgctcgagg | tcccgggcta | cgaggcggac | 360 |
| gacgtcctgg | ccagcctggc | caagaaggcg | gaaaaggagg | gctacgaggt | ccgcaccttc | 420 |
| accgccgaca | aagaccttta | ccagctcctt | tccgaccgca | tccacgtcct | ccaccccgag | 480 |
| gggtacctca | tcaccccggc | ctggccttgg | gaaaagtacg | gcctgaggcc | cgaccagtgg | 540 |
| gccgactacc | gggccctgac | cggggacgag | tccgacaacc | ttcccggggt | caagggcatc | 600 |
| ggggagaaga | cggcgaggaa | gcttctggag | gagtggggga | gcctggaagc | cctcctcaag | 660 |
| aacctggacc | ggctgaagcc | cgccatccgg | gagaagatcc | tggcccat | ggacgatctg | 720 |
| aagctctcct | gggacctggc | caaggtgcgc | accgacctgc | ccctggaggt | ggacttcgcc | 780 |
| aaaaggcggg | agcccgaccg | ggagaggctt | agggcctttc | tggagaggct | tgagtttggc | 840 |
| agcctcctcc | acgagttcgg | ccttctggaa | agccccaagg | ccctggagga | ggccccctgg | 900 |
| ccccgcgcgg | aaggggcctt | cgtgggcctt | gtgctttccc | gcaaggagcc | catgtggggc | 960 |
| gatcttctgg | ccctggccgc | cgccaggggc | ggccgcgtgc | accgggcagc | agaccccttg | 1020 |
| gcgggggctaa | aggacctcaa | ggaggtccgg | ggcctcctcg | ccaaggacct | cgccgtcttg | 1080 |
| gcctcgaggg | aggggctaga | cctcgtgccc | ggggacgacc | ccatgctcct | cgctacctc | 1140 |
| ctggaccctt | cgaacaccac | ccccgagggg | gtggcgcggc | gctacggggg | ggagtggacg | 1200 |
| gaggacgccg | cccaccgggc | cctcctctcg | gagaggctcc | atcggaacct | ccttaagcgc | 1260 |
| ctcgaggggg | aggagaagct | cctttggctc | taccacgagg | tggaaaagcc | cctctcccgg | 1320 |
| gtcctggccc | atatggaggc | cacgggggtg | cgcttgacg | tggcctatct | cagggccttg | 1380 |
| tccctggagc | ttgccgagga | gatcgccccg | ctcgaggccg | aggtcttccg | cctggccggc | 1440 |
| caccccttca | acctcaactc | ccgggaccag | ctggaaaagg | tcctctttga | cgagctaggg | 1500 |
| cttcccgcca | tcggcaagac | ggagaagacc | ggcaagcgct | ccaccagcgc | cgccgtcctg | 1560 |
| gaggccctcc | gcgaggccca | ccccatcg | gagaagatcc | tgcagtaccg | ggagctcacc | 1620 |
| aagctgaaga | gcacctacat | tgacccttg | ccggacctca | tccaccccag | gacgggccgc | 1680 |
| ctccacaccc | gcttcaacca | gacggccacg | gccacgggca | ggctaagtag | ctccgatccc | 1740 |
| aacctccaga | acatccccgt | ccgcaccccg | cttgggcaga | ggatccgccc | ggccttcac | 1800 |
| gccgaggagg | ggtggctatt | ggtggccctg | gactatagcc | agatagagct | cagggtgctg | 1860 |
| gcccacctct | ccggcgacga | gaacctgac | cggtcttcc | aggagggggc | ggacatccac | 1920 |
| acggagaccg | ccagctggat | gttcggcgtc | ccccgggagg | ccgtggaccc | cctgatgcgc | 1980 |

cccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc 960
 gatcttcttg ccctggccgc cgccaggggc ggccgcgtgc accgggcagc agaccccttg 1020
 gcggggctaa aggacctcaa ggaggtccgg ggctcctcg ccaaggacct cgccgtcttg 1080
 gcctcgaggg aggggctaga cctcgtgccc ggggacgacc ccatgctcct cgctacctc 1140
 ctggaccctt cgaacaccac ccccgagggg gtggcgcggc gctacggggg ggagtggacg 1200
 gaggacgccg cccaccgggc cctcctctcg gagaggctcc atcggaacct ccttaagcgc 1260
 ctcgaggggg aggagaagct cctttggctc taccacgagg tggaaaagcc cctctcccgg 1320
 gtcctggccc atatggaggc cacgggggtg cgctggacg tggcctatct cagggccttg 1380
 tccctggagg tggccgagga gatccgccgc ctcgaggccg aggtcttccg cctggccggc 1440
 cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
 cttcccgcca tcggcaagac ggagaagacc ggcaagcgt ccaccagcgc cgccgtcctg 1560
 gaggcctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
 aagctgaaga gcacctacat tgacccttg ccggacctca tccacccag gacgggccgc 1680
 ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
 aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggccttcac 1800
 gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
 gcccacctct ccggcgacga gaacctgatc cgggtcttcc aggagggggc ggacatccac 1920
 acggagaccg ccagctggat gttcggcgtc cccggggagg ccgtggacc cctgatgcgc 1980
 cgggcggcca agaccatcaa cttcggggtc ctctacggca tgcgggcca ccgcctctcc 2040
 caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
 gtggagaccc tcttcggccg ccgcgctac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgccc tccagggcac cgccgccgac 2280
 ctcatgaagc tggctatggt gaagctcttc ccaggctgg aggaaatggg ggccaggatg 2340
 ctcttcagg tcacaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
 gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggagggtg 2460
 gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 91

<211> 2517
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 91
 atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggctcctct ggtggacggc 60
 caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
 ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
 gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggccctacggg 240
 ggggtacaagg cgggcccgggc ccccacgccg gaggactttc cccggcaact cggcctcatc 300
 aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
 gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420
 accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccag 480
 ggggtacctca tcaccccggc ctggccttgg gaaaagtacg gcctgaggcc cgaccagtgg 540
 gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc 600
 ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
 aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg 720
 aagctctcct gggacctggc caaggtgccc accgacctgc ccctggaggt ggacttcgcc 780
 aaaaggcggg agcccgaccg ggagaggctt agggccttcc tggagaggct tgagtttggc 840
 agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
 ccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc 960
 gatcttctgg ccctggccgc cgccaggggc ggccgcgtgc accgggcagc agacccttg 1020
 gcgggggctaa aggacctcaa ggaggtccgg ggcctcctcg ccaaggacct cgccgtcttg 1080
 gcctcgaggg aggggctaga cctcgtgcc ggggacgacc ccatgctcct cgccctacctc 1140
 ctggaccctt cgaacaccac ccccgagggg gtggcgcggc gctacggggg ggagtggacg 1200
 gaggacgccg ccacccgggc cctcctctcg gagaggctcc atcggaacct ccttaagcgc 1260
 ctcgaggggg aggagaagct cctttggctc taccacgagg tggaaaagcc cctctcccgg 1320
 gtcttgggcc atatggaggc cacgggggtg cgcctggacg tggcctatct cagggccttg 1380
 tccttgagg tggccgagga gatcgccgc ctcgaggagg aggtcttccg cctggccggc 1440
 cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500

cttccccgcca tctggaagac ggagaagacc ggcaagcgct ccaccagcgc cgcggtcctg 1560
gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgacccttg ccggacctca tccaccccag gacggggcgc 1680
ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggcccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
gcccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggag ggacatccac 1920
acggagaccg ccagctggat gttcggcgtc ccccgaggag ccgtggaccc cctgatgcgc 1980
cgggcggcca agaccatcaa cttcggggtc ctctacggca tgcgggcca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgcgcgcgac 2280
ctcatgaagc tggctatggt gaagctcttc ccaggctgg aggaaatggg ggccaggatg 2340
ctccttcagg tccacaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
gagggtggga taggggagga ctggctctcc gccaaaggag accaccacca ccaccac 2517

<210> 92
<211> 2517
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 92
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggctctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
ccggtgcagg cggctctacg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggcccggc cccacgccg gaggacttcc cccggcaact cgccttcac 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360

| | |
|--|------|
| gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcacccctc | 420 |
| accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag | 480 |
| gggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg | 540 |
| gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc | 600 |
| ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag | 660 |
| aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg | 720 |
| aagctctcct gggacctggc caagggtgcg accgacctgc ccctggagggt ggacttcgcc | 780 |
| aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc | 840 |
| agcctcctcc acgagttcgg ctttctggaa agccccaagg ccctggagga ggccccctgg | 900 |
| cccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtggggc | 960 |
| gatcttcttg ccctggccgc cgccaggggc ggccgcgtgc accgggcagc agacccttg | 1020 |
| gcggggctaa aggacctcaa ggaggtccgg ggctcctcg ccaaggacct cgccgtcttg | 1080 |
| gcctcgaggg aggggctaga cctcgtgccc ggggacgacc ccatgctcct cgctacctc | 1140 |
| ctggaccctt cgaacaccac ccccgagggg gtggcgcggc gctacggggg ggagtggacg | 1200 |
| gaggacgccc cccaccgggc cctcctctcg gagaggctcc atcggaacct ccttaagcgc | 1260 |
| ctcgaggggg aggagaagct cctttggctc taccacgagg tggaaaagcc cctctcccgg | 1320 |
| gtcctggccc atatggaggc cacgggggtg cgcttgagc tggcctatct cagggccttg | 1380 |
| tccctggagg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc | 1440 |
| cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaagg | 1500 |
| cttcccgcca tcggcaagac ggagaagacc ggcaagcgct ccaccagcgc cgccgtcctg | 1560 |
| gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc | 1620 |
| aagctgaaga gcacctacat tgacccttg ccggacctca tccacccag gacgggcccgc | 1680 |
| ctccacacc gcttcaacca gacggccaag gccacgggca ggctaagtag ctccgatccc | 1740 |
| aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccc ggccttcac | 1800 |
| gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg | 1860 |
| gcccacctct ccggcgacga gaacctgac cgggctctcc aggaggggcg ggacatccac | 1920 |
| acggagaccg ccagctggat gttcggcgtc ccccgggagg ccgtggaccc cctgatgcgc | 1980 |
| cgggcccga agaccatcaa cttcggggtc ctctacggca tgtcggccca ccgcctctcc | 2040 |
| caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc | 2100 |

gatcttcttg ccctggccgc cgcaggggc ggccgcgtgc accgggcagc agacccttg 1020
gcggggctaa aggacctcaa ggaggtccgc ggctcctcg ccaaggacct cgccgtcttg 1080
gcctcgaggg aggggctaga cctcgtgcc gggaacgacc ccatgctcct cgccctacctc 1140
ctggaccctt cgaacaccac ccccgagggg gtggcgcggc gctacggggg ggagtggacg 1200
gaggacgccc cccaccgggc cctcctctcg gagaggtcc atcggaacct ccttaagcgc 1260
ctcgaggggg aggagaagct cctttggctc taccacgagg tggaaaagcc cctctcccgc 1320
gtcctggccc atatggaggg cacgggggtg cgctggacg tggcctatct cagggccttg 1380
tccctggagg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
cttcccgcca tcggcaagac gcaaaagacc ggcaagcgct ccaccagcgc cgccgtcctg 1560
gaggccctcc gcgaggccca ccccatcggt gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgacccttg ccggacctca tccaccccag gacgggcccgc 1680
ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
gcccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac 1920
acggagaccg ccagctggat gttcggcgtc ccccgaggagg ccgtggaccc cctgatgcgc 1980
cgggcgccca agaccatcaa cttcggggtc ctctacggca tgtcggccca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
gtggagacct tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
gtgcgggagg cgcccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
ctcatgaagc tggctatggt gaagctcttc ccaggtctgg aggaaatggg ggccaggatg 2340
ctccttcagg tccacaacga gctggctctc gaggccccaa aagagagggc ggaggccgtg 2400
gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 94
<211> 2517
<212> DNA
<213> Artificial

<220>

<223> Synthetic

<400> 94

| | |
|--|------|
| atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc | 60 |
| caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag | 120 |
| ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg | 180 |
| gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg | 240 |
| gggtacaagg cgggccgggc cccacagccg gaggactttc cccggcaact cgcctcctc | 300 |
| aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac | 360 |
| gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc | 420 |
| accgccgaca aagaccttta ccagctcctt tccgaccgca tcacgtcct ccaccccgag | 480 |
| gggtacctca tcaccccggc ctggctttgg gaaaagtacg gcctgaggcc cgaccagtgg | 540 |
| gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcctc | 600 |
| ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag | 660 |
| aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg | 720 |
| aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc | 780 |
| aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc | 840 |
| agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg | 900 |
| cccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc | 960 |
| gatcttctgg ccctggccgc cgccaggggc ggccgcgtgc accgggcagc agacccttg | 1020 |
| gcggggctaa aggacctcaa ggaggtccgg ggcctcctcg ccaaggacct cgccgtcttg | 1080 |
| gcctcgaggg aggggctaga cctcgtgccc ggggacgacc ccatgctcct cgctacctc | 1140 |
| ctggaccctt cgaacaccac ccccgagggg gtggcgcggc gctacggggg ggagtggacg | 1200 |
| gaggacgccg cccaccgggc cctcctctcg gagaggctcc atcggaacct ccttaagcgc | 1260 |
| ctcgaggggg aggagaagct cctttggctc taccacgagg tggaaaagcc cctctcccgg | 1320 |
| gtcctggccc atatggaggc cacgggggtg cgcttgagc tggcctatct cagggccttg | 1380 |
| tccttgagg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc | 1440 |
| cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg | 1500 |
| cttcccgcca tcggcaagac ggagaagacc ggcaagcgct ccaccagcgc cgccgtcctg | 1560 |

gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagcaccg ggagctcacc 1620
aagctgaaga gcacctacat tgacccttg ccggacctca tccacccag gacgggcccgc 1680
ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct caggggtgctg 1860
gcccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggag ggacatccac 1920
acggagaccg ccagctggat gttcggcgtc ccccgggagg ccgtggaccc cctgatgcgc 1980
cgggcggcca agaccatcaa cttcggggtc ctctacggca tgcggccca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgcgcgcgac 2280
ctcatgaagc tggctatggt gaagctcttc ccaggctgg aggaaatggg ggccaggatg 2340
ctccttcagg tccacaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggagggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 95
<211> 2517
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 95
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
ccggtgcagg cggctctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggctct tgacgccaag gcccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggcccggc cccacgcgc gaggacttct cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480

| | | | | | | |
|------------|------------|------------|------------|-------------|-------------|------|
| gggtaccta | tcaccccg | ctggctttg | gaaaagtac | gcctgaggcc | cgaccagtgg | 540 |
| gccgactacc | gggccctgac | cggggacgag | tccgacaacc | ttcccgggggt | caagggcatc | 600 |
| ggggagaaga | cggcgaggaa | gcttctggag | gagtggggga | gcctggaagc | cctcctcaag | 660 |
| aacctggacc | ggctgaagcc | cgccatccgg | gagaagatcc | tggcccacat | ggacgatctg | 720 |
| aagctctcct | gggacctggc | caaggtgctc | accgacctgc | ccctggagggt | ggacttcgcc | 780 |
| aaaaggcggg | agcccgaccg | ggagaggctt | agggcctttc | tggagaggct | tgagtttggc | 840 |
| agcctcctcc | acgagttcgg | ccttctggaa | agccccaagg | ccctggagga | ggccccctgg | 900 |
| cccccgccgg | aaggggcctt | cgtgggcttt | gtgctttccc | gcaaggagcc | catgtggggc | 960 |
| gatcttcttg | ccctggccgc | cgccaggggc | ggccgcgtgc | accgggcagc | agacccttg | 1020 |
| gcggggctaa | aggacctcaa | ggaggtccgg | ggcctcctcg | ccaaggacct | cgccgtcttg | 1080 |
| gcctcgaggg | aggggctaga | cctcgtgccc | ggggacgacc | ccatgctcct | cgcctacctc | 1140 |
| ctggaccctt | cgaacaccac | ccccgagggg | gtggcgcggc | gctacggggg | ggagtggacg | 1200 |
| gaggacgccg | cccaccgggc | cctcctctcg | gagaggctcc | atcggaacct | ccttaagcgc | 1260 |
| ctcgaggggg | aggagaagct | cctttggctc | taccacgagg | tggaaaagcc | cctctcccgg | 1320 |
| gtcctggccc | atatggaggc | cacgggggtg | cgctggacg | tggcctatct | cagggccttg | 1380 |
| tccctggagg | tggccgagga | gatcgcccgc | ctcgaggccg | aggtcttccg | cctggccggc | 1440 |
| cacccttca | acctcaactc | cggggaccag | ctggaaaggg | tcctctttga | cgagctaggg | 1500 |
| cttcccgcca | tcggcaagac | ggagaagacc | ggcaagcgct | ccaccagcgc | cgccgtcctg | 1560 |
| gaggccctcc | gcgaggccca | ccccatcgtg | gagaagatcc | tgcagtaccg | ggagctcacc | 1620 |
| aagctgaaga | acacctacat | tgaccccttg | cgggacctca | tccaccccag | gacgggcccgc | 1680 |
| ctccacaccc | gcttcaacca | gacggccacg | gccacgggca | ggctaagtag | ctccgatccc | 1740 |
| aacctccaga | acatccccgt | ccgcaccccg | cttgggcaga | ggatccgccg | ggccttcac | 1800 |
| gccgaggagg | ggtggctatt | ggtggccctg | gactatagcc | agatagagct | caggggtgctg | 1860 |
| gcccacctct | cgggcgacga | gaacctgac | cgggtcttcc | aggagggggc | ggacatccac | 1920 |
| acggagaccg | ccagctggat | gttcggcgct | ccccgggagg | ccgtggaccc | cctgatgcgc | 1980 |
| cgggcggcca | agaccatcaa | cttcggggct | ctctacggca | tgctggccca | ccgcctctcc | 2040 |
| caggagctag | ccatccctta | cgaggaggcc | caggccttca | ttgagcgcta | ctttcagagc | 2100 |
| ttccccaagg | tgcgggcctg | gattgagaag | accctggagg | agggcaggag | gcgggggtac | 2160 |

gtggagaccc tcttcggccg ccgcccgtac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cgcccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
 ctcatgaagc tggctatggt gaagctcttc cccaggctgg aggaaatggg ggccaggatg 2340
 ctcttccagg tccacaacga gctggctctc gagggcccaa aagagagggc ggaggccgtg 2400
 gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
 gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 96
 <211> 2517
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 96
 atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
 caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
 ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
 gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
 ggggtacaagg cgggcccggc cccacgcgcg gaggactttc cccggcaact cgccctcatc 300
 aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
 gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420
 accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccag 480
 gggtagctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
 gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcac 600
 ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
 aacctggacc ggctgaagcc cgccatcccg gagaagatcc tggccacat ggacgatctg 720
 aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780
 aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
 agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
 cccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc 960
 gatcttctgg ccctggccgc cgccaggggc ggccgcgtgc accgggcagc agacccttg 1020
 gcgggggctaa aggacctcaa ggaggtccgg ggccctcctcg ccaaggacct cgccgtcttg 1080

gcctcgaggg aggggctaga cctcgtgccc ggggacgacc ccatgctcct cgcctacctc 1140
ctggaccett cgaacaccac ccccgagggg gtggcgcggc gctacggggg ggagtggacg 1200
gaggacgccc cccaccgggc cctcctctcg gagaggctcc atcggaacct ccttaagcgc 1260
ctcgaggggg aggagaagct cctttggctc taccacgagg tggaaaagcc cctctcccgg 1320
gtcctggccc atatggaggc cacgggggtg cgcctggacg tggcctatct cagggccttg 1380
tccctggagg tggccgagga gatcgcccg ctcgaggccg aggtcttccg cctggccggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
cttcccgcga tcggcaagac ggagaagacc ggcaagcgct ccaccagcgc cgccgtcctg 1560
gaggccctcc gcgaggccca ccccatcggt gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacgt ggacccttg ccggacctca tccaccccag gacgggcccgc 1680
ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
gcccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggag ggacatccac 1920
acggagaccg ccagctggat gttcggcgtc ccccgaggag ccgtggaccc cctgatgcgc 1980
cgggcggcca agaccatcaa cttcggggtc ctctacggca tgtcggccca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
ctcatgaagc tggctatggt gaagctcttc ccagggtgg aggaaatggg ggccaggatg 2340
ctccttcagg tccacaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 97
<211> 2517
<212> DNA
<213> Artificial

<220>
<223> Synthetic

| | |
|--|------|
| <400> 97 | |
| atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc | 60 |
| caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag cgggggggag | 120 |
| ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg | 180 |
| gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg | 240 |
| gggtacaagg cgggccgggc cccacgccg gaggactttc cccggcaact cgccctcatc | 300 |
| aaggagctgg tggacctcct ggggctggcg cgctcgagg tcccgggcta cgaggcggac | 360 |
| gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc | 420 |
| accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag | 480 |
| gggtacctca tcaccccggc ctggctttgg gaaaagtacg gcctgaggcc cgaccagtgg | 540 |
| gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc | 600 |
| ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag | 660 |
| aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg | 720 |
| aagctctcct gggacctggc caaggtgcmc accgacctgc ccctggagggt ggacttcgcc | 780 |
| aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc | 840 |
| agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg | 900 |
| cccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtggggc | 960 |
| gatcttctgg ccctggccgc cgccaggggc ggccgcgtgc accgggcagc agaccccttg | 1020 |
| gcgggggctaa aggacctcaa ggaggtccgg ggcctcctcg ccaaggacct cgccgtcttg | 1080 |
| gcctcgaggg aggggctaga cctcgtgccc ggggacgacc ccatgctcct cgcctacctc | 1140 |
| ctggaccctt cgaacaccac ccccgagggg gtggcgcggc gctacggggg ggagtggacg | 1200 |
| gaggacgccg cccaccgggc cctcctctcg gagaggctcc atcggaacct ccttaagcgc | 1260 |
| ctcgaggggg aggagaagct cctttggctc taccacgagg tggaaaagcc cctctcccgg | 1320 |
| gtcctggccc atatggaggc cacgggggtg cgctggacg tggcctatct cagggccttg | 1380 |
| tccctggagg tggccgagga gatcgccgcg ctcgaggccg aggtcttccg cctggccggc | 1440 |
| caccccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg | 1500 |
| cttcccgcca tcggcaagac ggagaagacc ggcaagcgt ccaccagcgc cgccgtcctg | 1560 |
| gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc | 1620 |
| aagctgaaga gcacctacat tgaccccttg ccgagcctcg tccacccag gacgggccgc | 1680 |

ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct caggggtgctg 1860
gcccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggag ggacatccac 1920
acggagaccg ccagctggat gttcggcgtc ccccgggagg ccgtggaccc cctgatgcgc 1980
cgggcggcca agaccatcaa cttcggggtc ctctacggca tgcgggcca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
tcccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
ctcatgaagc tggctatggt gaagctcttc cccaggctgg aggaaatggg ggccaggatg 2340
ctccttcagg tccacaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggagggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggag accaccacca ccaccac 2517

<210> 98
<211> 2517
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 98
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
ccggtgcagg cggctctacg cttcgccaag agcctcctca aggcctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gcccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggccgggc cccacgcgcg gaggacttct cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggctggcg cgctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccgcc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540

| | | | | | | |
|------------|-------------|------------|------------|-------------|-------------|------|
| gccgactacc | gggccctgac | cggggacgag | tccgacaacc | ttcccgggggt | caagggcatc | 600 |
| ggggagaaga | cggcgaggaa | gcttctggag | gagtggggga | gcctggaagc | cctcctcaag | 660 |
| aacctggacc | ggctgaagcc | cgccatccgg | gagaagatcc | tggcccacat | ggacgatctg | 720 |
| aagctctcct | gggacctggc | caaggtgcgc | accgacctgc | ccctggagggt | ggacttcgcc | 780 |
| aaaaggcggg | agccccgaccg | ggagaggctt | agggcctttc | tggagaggct | tgagtttggc | 840 |
| agcctcctcc | acgagttcgg | ccttctggaa | agccccaagg | ccctggagga | ggccccctgg | 900 |
| ccccgcgagg | aaggggcctt | cgtgggcctt | gtgctttccc | gcaaggagcc | catgtggggc | 960 |
| gatcttctgg | ccctggccgc | cgccaggggc | ggccgcgtcc | accggggccc | cgagccttat | 1020 |
| aaagccctca | gggacctgaa | ggaggcgagg | gggcttctcg | caaagacct | gagcgttctg | 1080 |
| gccctgaggg | aaggccttgg | cctcccgccc | ggcgacgacc | ccatgctcct | cgcctacctc | 1140 |
| ctggaccctt | cgaacaccac | ccccgagggg | gtggcccgcc | gctacggcgg | ggagtggacg | 1200 |
| gaggaggcgg | gggagcgggc | cgccctttcc | gagaggctct | tcgccaacct | gcttaagagg | 1260 |
| cttgaggggg | aggagaggct | cctttggcct | taccgggagg | tggagaggcc | cctttccgct | 1320 |
| gtcctggccc | atatggaggc | cacgggggtg | cgctggacg | tggcctatct | cagggccttg | 1380 |
| tccctggagg | tggccgagga | gatcgccgcg | ctcgaggccg | aggtcttccg | cctggccggc | 1440 |
| cacccttca | acctcaactc | ccgggaccag | ctggaaaggg | tcctctttga | cgagctaggg | 1500 |
| cttcccgcca | tcggcaagac | gcaaaagacc | ggcaagcgct | ccaccagcgc | cgcgctcctg | 1560 |
| gaggccctcc | gcgaggccca | ccccatcggt | gagaagatcc | tgagtagacc | ggagctcacc | 1620 |
| aagctgaaga | gcacctacat | tgacctcttg | ccggacctca | tccaccccag | gacgggcccgc | 1680 |
| ctccacaccc | gcttcaacca | gacggccacg | gccacgggca | ggctaagtag | ctccgatccc | 1740 |
| aacctccaga | acatccccgt | ccgcaccccg | cttgggcaga | ggatccgcgc | ggccttcctc | 1800 |
| gccgaggagg | ggtggctatt | ggtggccctg | gactatagcc | agatagagct | caggggtgctg | 1860 |
| gccacctct | ccggcgacga | gaacctgac | cggtcttccc | aggagggggc | ggacatccac | 1920 |
| acggagaccg | ccagctggat | gttcggcgct | ccccgggagg | ccgtggaccc | cctgatgcgc | 1980 |
| cgggcgggca | agaccatcaa | cttcgggggt | ctctacggca | tgtcggccca | cgcctctccc | 2040 |
| caggagctag | ccatccctta | cgaggaggcc | caggccttca | ttgagcgcta | ctttcagagc | 2100 |
| ttccccaagg | tgcgggcctg | gattgagaag | accctggagg | agggcaggag | gcgggggtac | 2160 |
| gtggagaccc | tcttcggccg | ccgcgcgtac | gtgccagacc | tagaggcccc | ggtgaagagc | 2220 |
| gtgcgggagg | cggccgagcg | catggccttc | aacatgcccg | tccagggcac | cgcgcgcgac | 2280 |

ctcatgaagc tggctatggt gaagctcttc cccaggctgg aggaaatggg ggccaggatg 2340
ctccttcagg tccacaacga gctggctctc gaggccccaa aagagagggc ggaggccgtg 2400
gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 99
<211> 2517
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 99
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
ccggtgcagg cggctctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggccgggc ccccacgccg gaggactttc cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccggc ctggcctttg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggc caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg 720
aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
agcctcctcc acgagttcgg ctttctggaa agccccaagg ccctggagga ggccccctgg 900
ccccgcggg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggccc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg 1080
gccctgaggg aaggccttgg cctcccggcc ggcgacgacc ccctgctcct cgcctacctc 1140

ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200
gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttgggagg 1260
cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct 1320
gtcctggccc atatggaggc cacgggggtg cgcttgacg tggcctatct cagggccttg 1380
tccctggagg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
cttcccgcga tcggcaagac gcaaaagacc ggcaagcgt ccaccagcg cgccgtcctg 1560
gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgaccccttg ccggacctca tccacccag gacgggcccgc 1680
ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
gccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac 1920
acggagaccg ccagctggat gttcggcgtc ccccgaggagg ccgtggacc cctgatgcgc 1980
cgggcggcca agaccatcaa cttcggggtc ctctacggca tgcggccca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
gtggagacc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
ctcatgaagc tggctatggt gaagctcttc ccaggctgg aggaaatggg ggccaggatg 2340
ctccttcagg tccacaacga gctggtcctc gaggcccaa aagagagggc ggaggccgtg 2400
gcccggtg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggagggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 100
<211> 2517
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 100
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60

| | |
|--|------|
| caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag | 120 |
| ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg | 180 |
| gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg | 240 |
| gggtacaagg cgggcccggc cccacgccg gaggactttc cccggcaact cgccctcatc | 300 |
| aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac | 360 |
| gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc | 420 |
| accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccag | 480 |
| gggtacctca tcaccccggc ctggccttgg gaaaagtacg gcctgaggcc cgaccagtgg | 540 |
| gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc | 600 |
| ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag | 660 |
| aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg | 720 |
| aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggagggt ggacttcgcc | 780 |
| aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc | 840 |
| agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg | 900 |
| ccccgcgcgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc | 960 |
| gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat | 1020 |
| aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg | 1080 |
| gccctgaggg aaggccttgg cctcccgcgc ggcgacgacc ccatgctcct cgcctacctc | 1140 |
| ctggaccctt cgaacaccac ccccaggggg gtggcccggc gctacggcgg ggagtggacg | 1200 |
| gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gtggaagagg | 1260 |
| cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct | 1320 |
| gtcctggccc atatggaggc cacgggggtg cgctggacg tggcctatct cagggccttg | 1380 |
| tccttgaggg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc | 1440 |
| cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg | 1500 |
| cttcccgcga tcggcaagac gcaaaagacc ggcaagcgt ccaccagcgc cgccgtcctg | 1560 |
| gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc | 1620 |
| aagctgaaga gcacctacat tgacccttg ccggacctca tccaccccag gacgggccgc | 1680 |
| ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc | 1740 |

aacctccaga acatccccgt ccgcacccccg cttgggcaga ggatccgccg ggccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct caggggtgctg 1860
gccacctct ccggcgacga gaacctgac cgggtcttcc aggaggggag ggacatccac 1920
acggagaccg ccagctggat gttcggcgtc ccccgggagg ccgtggaccc cctgatgcgc 1980
cgggcggcca agaccatcaa cttcggggtc ctctacggca tgtcggccca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccc ggtgaagagc 2220
gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgcgcgcgac 2280
ctcatgaagc tggctatggt gaagctcttc ccaggctgg aggaaatggg ggccaggatg 2340
ctccttcagg tccacaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 101
<211> 2526
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 101
atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
ggccaccacc tggcctaccg caccttcttc gcctgaagg gcctcaccac gagccggggc 120
gaaccggtgc aggcggtcta cggttcgcc aagagcctcc tcaaggccct gaaggaggac 180
gggtacaagg ccgtcttcgt ggtctttgac gccaaaggccc cctccttcg ccacgaggcc 240
tacgaggcct acaaggcggg gagggccccg acccccgagg acttcccccg gcagctcgcc 300
ctcatcaagg agctggtgga cctcctgggg tttaccgcc tcgaggtccc cggctacgag 360
gcggacgacg ttctcgccac cctggccaag aaggcggaag aggggggta cgaggtgcgc 420
atcctcaccg ccgaccgga cctctaccaa ctcgctctcc accgcgtcgc cgtcctccac 480
cccgagggcc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag 540
cagtgggtgg acttccgcgc cctcgtgggg gaccctccg acaacctccc cggggtaag 600
ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt gggaagcct ggaaaacctc 660

| | |
|--|------|
| ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg | 720 |
| gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctccc cctggagggtg | 780 |
| gacctcgccc agggggcgga gcccgaccgg gaggggctta gggccttcct ggagaggctg | 840 |
| gagttcggca gcctcctcca cgagttcggc ctcttgagg cccccgcccc cctggaggag | 900 |
| gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctcccg ccccgagccc | 960 |
| atgtgggagg agcttaaagc cctggccgcc tgcaggagcg gccgggtgca ccgggcagca | 1020 |
| gaccccttgg cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc | 1080 |
| gccgtcttgg cctcgaggga ggggctagac ctctgtcccg gggacgacct catgctctc | 1140 |
| gcctacctcc tggacccttc gaacaccacc cccgaggggg tggcgcgggc ctacgggggg | 1200 |
| gagtggagcg aggacgccc ccaccgggccc ctctctcgg agaggctcca tcggaacctc | 1260 |
| cttaagcgcc tcgaggggga ggagaagctc ctttggtct accacgaggt ggaaaagccc | 1320 |
| ctctcccggg tcctggccca tatggaggcc accggggtag gccgggacgt ggcctacctt | 1380 |
| caggcccttt ccctggagct tgccggagg atccgccgcc tcgaggagga ggtcttccgc | 1440 |
| ttggcgggcc accccttcaa cctcaactcc cgggaccagc tggaaagggt gctctttgac | 1500 |
| gagcttaggc ttccgcctt ggggaagacg caaaagacag gcaagcgctc caccagcgcc | 1560 |
| gcggtgctgg aggccctacg ggaggccac cccatcgtgg agaagatcct ccagcaccgg | 1620 |
| gagctacca agctcaagaa cacctacgtg gacccctcc caagcctcgt ccacccgagg | 1680 |
| acgggccgcc tccacaccg cttcaaccag acggccacgg ccacggggag gcttagtagc | 1740 |
| tccgaccca acctgcagaa catccccgtc cgcacccct tgggcccagag gatccgccgg | 1800 |
| gccttcgtgg ccgaggcggg ttgggcgttg gtggccctgg actatagcca gatagagctc | 1860 |
| cgcgtcctcg cccacctctc cggggacgaa aacctgatca gggctctcca ggaggggaag | 1920 |
| gacatcgcca ccagaccgc aagctggatg ttggcgctcc ccccgaggc cgtggacccc | 1980 |
| ctgatcgcc gggcgccaa gacggtgaac ttggcgctcc tctacggcat gtccgccat | 2040 |
| aggctctccc aggagcttgc catcccctac gaggaggcgg tggcctttat agagcgctac | 2100 |
| ttccaaagct tcccaaaggt gcgggcctgg atagaaaaga ccctggagga ggggaggaag | 2160 |
| cggggctacg tggaaaccct cttcggaaga aggcgctacg tgcccgacct caacgcccgg | 2220 |
| gtgaagagcg tcaggaggc cgcggagcgc atggccttca acatgcccgt ccagggcacc | 2280 |
| gccgccgacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg | 2340 |

gcccgcatgc tctccaggt ccacaacgag ctctcctgg agggccccc agcgcgggcc 2400
 gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatccctt cgccgtgccc 2460
 ctggaggtgg aggtggggat gggggaggac tggctttccg ccaagggta ccaccaccac 2520
 caccac 2526

<210> 102
 <211> 2526
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 102
 atgaattccg aggcgatgct tccgctcttt gaaccctaaag gccgggtcct cctggtggac 60
 ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
 gaaccggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac 180
 ggggtacaagg ccgtcttcgt ggtctttgac gccaaaggccc cctccttccg ccacgaggcc 240
 tacgaggcct acaaggcggg gagggccccc acccccgagg acttcccccg gcagctcgcc 300
 ctcatcaagg agctggtgga cctcctgggg ttaccgccg tcgagggtccc cggctacgag 360
 gcggacgacg ttctcgccac cctggccaag aaggcggaag aggaggggta cgaggtgcgc 420
 atcctcaccg ccgaccgcga cctctacca ctcgtctccg accgcgtcgc cgtcctccac 480
 cccgagggcc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag 540
 cagtgggtgg acttccgcgc cctcgtgggg gacccctccg acaacctccc cgggggtcaag 600
 ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc 660
 ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg 720
 gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctccc cctggagggtg 780
 gacctcgccc aggggcggga gcccgaccgg gaggggctta gggccttcct ggagaggctg 840
 gagttcggca gcctcctcca cgagttcggc ctcttgagg ccccgcccc cctggaggag 900
 gcccctggc ccccgccgga aggggccttc gtgggcttcg tctctcccg ccccgagccc 960
 atgtgggcgg agcttaaagc cctggccgcc tgcaggggag gccgcgtgca ccgggcagca 1020
 gaccccttgg cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc 1080
 gccgtcttgg cctcgaggga ggggctagac ctctgcccc gggacgaccc catgctcctc 1140
 gcctacctcc tggaccttc gaacaccacc cccgaggggg tggcgggcg ctacgggggg 1200

| | | | | | | |
|-------------|------------|-------------|-------------|-------------|------------|------|
| gagtggacgg | aggacgccgc | ccaccggggcc | ctcctctcgg | agaggctcca | tcggaacctc | 1260 |
| cttaagcgcc | tcgaggggga | ggagaagctc | ctttggctct | accacgaggt | ggaaaagccc | 1320 |
| ctctcccggg | tcctggccca | tatggaggcc | accgggggtac | ggcgggacgt | ggcctacctt | 1380 |
| caggcccttt | ccctggagct | tgcggaggag | atccgccgcc | tcgaggagga | ggtcttcgc | 1440 |
| ttggcgggcc | accccttcaa | cctcaactcc | cgggaccagc | tggaaagggg | gctctttgac | 1500 |
| gagcttaggc | ttcccgctt | ggggaagacg | caaaagacag | gcaagcgctc | caccagcgcc | 1560 |
| gcggtgctgg | aggccctacg | ggaggcccac | cccatcgtgg | agaagatcct | ccagcaccgg | 1620 |
| gagctcacca | agctcaagaa | cacctacgtg | gacccccctc | caagcctcgt | ccacccgagg | 1680 |
| acgggcccgc | tccacacccg | cttcaaccag | acggccacgg | ccacggggag | gcttagtagc | 1740 |
| tccgacccca | acctgcagaa | catccccgtc | cgcacccccct | tggggccagag | gatccgccgg | 1800 |
| gccttcgtgg | ccgaggcggg | ttgggcgttg | gtggccctgg | actatagcca | gatagagctc | 1860 |
| cgcgctctcg | cccacctctc | cggggacgaa | aacctgatca | gggtcttcca | ggaggggaag | 1920 |
| gacatccaca | cccagaccgc | aagctggatg | ttcggcgctc | ccccggaggc | cgtggacccc | 1980 |
| ctgatgcgcc | gggcggccaa | gacggtgaac | ttcggcgctc | tctacggcat | gtccgcccac | 2040 |
| aggctctccc | aggagcttgc | catcccctac | gaggaggcgg | tggcctttat | agagcgctac | 2100 |
| ttccaaagct | tccccaaggt | gcgggcctgg | atagaaaaga | ccctggagga | ggggaggaag | 2160 |
| cggggctacg | tggaaaccct | cttcggaaga | aggcgctacg | tgcccgacct | caacgcccgg | 2220 |
| gtgaagagcg | tcagggaggc | cgcggaggcc | atggccttca | acatgcccgt | ccagggcacc | 2280 |
| gccgccgacc | tcatgaagct | cgccatggtg | aagctcttcc | ccgcctccg | ggagatgggg | 2340 |
| gcccgcgatgc | tcctccaggt | ccacaacgag | ctcctcctgg | aggcccccca | agcgcgggcc | 2400 |
| gaggaggtgg | cggctttggc | caaggaggcc | atggagaagg | cctatcccct | cgccgtgccc | 2460 |
| ctggaggtgg | aggtggggat | gggggaggac | tggctttccg | ccaagggtca | ccaccaccac | 2520 |
| caccac | | | | | | 2526 |

| | |
|-------|------------|
| <210> | 103 |
| <211> | 2526 |
| <212> | DNA |
| <213> | Artificial |
| <220> | |
| <223> | Synthetic |
| <400> | 103 |

| | |
|---|------|
| atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctgggtggac | 60 |
| ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc | 120 |
| gaaccggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac | 180 |
| gggtacaagg ccgtcttcgt ggtctttgac gccaaagggc cctccttcg ccacgaggcc | 240 |
| tacgaggcct acaaggcggg gagggccccg acccccgagg acttccccg gcagctcgcc | 300 |
| ctcatcaagg agctggtgga cctcctgggg ttaccgccg tcgagggtccc cggctacgag | 360 |
| gcggaacgacg ttctcgccac cctggccaag aaggcggaaa aggaggggta cgagggtgcgc | 420 |
| atcctcaccg ccgaccgcga cctctaccaa ctggtctccg accggtcgcc cgtcctccac | 480 |
| ccccgaggcc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag | 540 |
| cagtgggtgg acttccgcgc cctcgtgggg gacccctccg acaacctccc cgggggtcaag | 600 |
| ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc | 660 |
| ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg | 720 |
| gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctccc cctggagggtg | 780 |
| gacctcgccc aggggcggga gcccgaccgg gaggggctta gggccttcct ggagaggctg | 840 |
| gagttcggca gcctcctcca cgagttcggc ctcttgagg cccccgcccc cctggaggag | 900 |
| gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctcccg ccccgagccc | 960 |
| atgtgggcgg agcttaaagc cctggccgcc tgcaggggag gccgcgtgca ccgggcagca | 1020 |
| gaccccttgg cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc | 1080 |
| gccgtcttgg cctcgaggga ggggctagac ctctgccccg gggacgaccc catgctcctc | 1140 |
| gcctacctcc tggacccttc gaacaccacc cccgaggggg tggcgcggcg ctacgggggg | 1200 |
| gagtggacgg aggacgcgc ccaccgggccc ctctctcgg agaggctcca tcggaacctc | 1260 |
| cttaagcgcc tcgaggggga ggagaagctc ctttggctct accacgaggt ggaaaagccc | 1320 |
| ctctccccgg tcctggccca tatggaggcc accggggtag ggcgggacgt ggcctacctt | 1380 |
| caggcccttt ccctggagct tgcggaggag atccgccgcc tcgaggagga ggtcttccgc | 1440 |
| ttggcggggc accccttcaa cctcaactcc cgggaccagc tggaaagggt gctctttgac | 1500 |
| gagcttaggc ttcccgctt ggggaagacg caaaagacag gcaagcgctc caccagcgcc | 1560 |
| gcggtgctgg aggcctacg ggaggccac cccatcgtgg agaagatcct ccagcaccgg | 1620 |
| gagctacca agctcaagaa cacctacgtg gacccctcc caagcctcgt ccacccgagg | 1680 |
| acgggccgcc tcacacccg cttcaaccag acggccacgg ccacggggag gcttagtagc | 1740 |

| | | | | | | |
|------------|------------|-------------|------------|------------|------------|------|
| tccgacccca | acctgcagaa | catccccgtc | cgcaccccct | tgggccagag | gatccgcccg | 1800 |
| gccttcgttg | ccgaggcggg | ttgggcgttg | gtggcccttg | actatagcca | gatagagctc | 1860 |
| cgcgtcctcg | cccacctctc | cggggacgaa | aacctgatca | gggtcttcca | ggaggggaag | 1920 |
| gacatccaca | cccagaccgc | aagctggatg | ttcggcgtcc | ccccggaggc | cgtggacccc | 1980 |
| ctgatgcgcc | gggcggccaa | gacgggtgaac | ttcggcgtcc | tctacggcat | gtccgcccac | 2040 |
| aggctctccc | aggagcttgc | catcccctac | gaggaggcgg | tggcctttat | agagcgctac | 2100 |
| ttccaaagct | tccccaaggt | gcgggccttg | atagaaaaga | ccctggagga | ggggaggaag | 2160 |
| cggggctacg | tggaaaccct | cttcggaaga | aggcgctacg | tgcccgacct | caacgcccgg | 2220 |
| gtgaagagcg | tcagggaggc | cgcggagcgc | atggccttca | acatgcccgt | ccagggcacc | 2280 |
| gccgccgacc | tcatgaagct | cgccatggtg | aagctcttcc | ccgcctccg | ggagatgggg | 2340 |
| gcccgcacgc | tcctccaggt | cgccaacgag | ctcctcctgg | agggccccca | agcgcgggcc | 2400 |
| gaggaggtgg | cggctttggc | caaggaggcc | atggagaagg | cctatcccct | cgccgtgccc | 2460 |
| ctggaggtgg | aggtggggat | gggggaggac | tggctttccg | ccaagggtca | ccaccaccac | 2520 |
| caccac | | | | | | 2526 |

| | |
|-------|------------|
| <210> | 104 |
| <211> | 2526 |
| <212> | DNA |
| <213> | Artificial |

<220>
<223> Synthetic

| | | | | | | | |
|------------|------------|------------|-------------|------------|------------|--|-----|
| <400> | 104 | | | | | | |
| atgaattccg | aggcgatgct | tccgctcttt | gaacccaaag | gccgggtcct | cctggtggac | | 60 |
| ggccaccacc | tggcctaccg | caccttcttc | gccctgaagg | gcctcaccac | gagccggggc | | 120 |
| gaaccggtgc | aggcgtcta | cggcttcgcc | aagagcctcc | tcaaggccct | gaaggaggac | | 180 |
| gggtacaagg | ccgtcttcgt | ggtctttgac | gccaaagccc | cctccttcgc | ccacgaggcc | | 240 |
| tacgaggcct | acaaggcggg | gagggccccg | acccccgagg | acttcccccg | gcagctcgcc | | 300 |
| ctcatcaagg | agctggtgga | cctcctgggg | tttaccgcgc | tcgaggctcc | cggctacgag | | 360 |
| gcggacgacg | ttctcgccac | cctggccaag | aaggcggaaa | aggaggggta | cgaggtgcgc | | 420 |
| atcctcaccg | ccgaccgcga | cctctaccaa | ctcgtctccg | accgcgtcgc | cgctctccac | | 480 |
| cccgagggcc | acctcatcac | cccgagtggt | ctttggggaga | agtacggcct | caggccggag | | 540 |

| | |
|--|------|
| cagtgggtgg acttccgcgc cctcgtgggg gacccctccg acaacctccc cggggtcaag | 600 |
| ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc | 660 |
| ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg | 720 |
| gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctccc cctggagggtg | 780 |
| gacctcggcc aggggcggga gcccgaaccg gaggggctta gggccttcct ggagaggctg | 840 |
| gagttcggca gcctcctcca cgagttcggc ctcttgagg ccccgcccc cctggaggag | 900 |
| gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctccc ccccgagccc | 960 |
| atgtgggagg agcttaaagc cctggccgcc tgcaggggag gccgcgtgca cggggcagca | 1020 |
| gaccccttgg cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc | 1080 |
| gccgtcttgg cctcgaggga ggggctagac ctctgtcccg gggacgaccc catgctcctc | 1140 |
| gcctacctcc tggacccttc gaacaccacc cccgaggggg tggcgcgagg ctacgggggg | 1200 |
| gagtggacgg aggacggcgc ccaccgggccc ctctctcgg agaggctcca tcggaacctc | 1260 |
| cttaagcgcc tcgaggggga ggagaagctc ctttggctct accacgaggt ggaaaagccc | 1320 |
| ctctcccggg tcctggccca tatggaggcc accggggtac ggcgggacgt ggccctacct | 1380 |
| caggcccttt ccctggagct tgcggaggag atccgccgcc tcgaggagga ggtcttcgcg | 1440 |
| ttggcgggccc accccttcaa cctcaactcc cgggaccagc tggaaagggt gctctttgac | 1500 |
| gagcttaggc ttcccgctt gaagaagacg aagaagacag gcaagcgctc caccagcgcc | 1560 |
| gcggtgctgg agggcctacg ggaggccac cccatcgtgg agaagatcct ccagcaccgg | 1620 |
| gagctcacca agctcaagaa cacctacgtg gacccctcc caagcctcgt ccacccgagg | 1680 |
| acgggcggcc tccacacccg cttcaaccag acggccacgg ccacggggag gcttagtagc | 1740 |
| tccgacccca acctgcagaa catccccgtc cgcacccct tgggcccagag gatccggcgg | 1800 |
| gccttcgtgg ccgaggcggg ttgggcgttg gtggccctgg actatagcca gatagagctc | 1860 |
| cgcgtcctcg cccacctctc cggggacgaa aacctgatca gggctctcca ggaggggaag | 1920 |
| gacatccaca cccagaccgc aagctggatg ttcggcgctc ccccgaggc cgtggacccc | 1980 |
| ctgatgcgcc gggcggccaa gacgggtgaac ttcggcgctc tctacggcat gtccgccc | 2040 |
| aggctctccc aggagcttgc catccccctac gaggaggcgg tggcctttat agagcgctac | 2100 |
| ttccaaagct tccccagggt gcgggccttg atagaaaaga ccctggagga ggggaggaag | 2160 |
| cggggctacg tggaaacctt cttcggaaga aggcgctacg tgcccgacct caacgcccgg | 2220 |
| gtgaagagcg tcagggaggc cgcggagcgc atggccttca acatgcccgt ccagggcacc | 2280 |

gcccgcgacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg 2340
 gcccgcagtc tctccaggt cgccaacgag ctctctctgg agggccccc aagcgggggc 2400
 gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatcccct cgccgtgccc 2460
 ctggaggtgg aggtggggat gggggaggac tggctttccg ccaagggta ccaccaccac 2520
 caccac 2526

<210> 105
 <211> 2517
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 105
 atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
 caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
 ccggtgcagg cggctctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
 gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
 ggggtacaagg cgggcccggc ccccacgccg gaggactttc cccggcaact cgccctcatc 300
 aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
 gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420
 accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
 ggggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
 gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggc caagggcatc 600
 ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
 aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccatat ggacgatctg 720
 aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780
 aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
 agcctcctcc acgagttcgg cttcttgaa agccccaagg ccctggagga gggccctgg 900
 cccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc 960
 gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
 aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg 1080

```

gccctgaggg aaggccttgg cctcccgccc ggcgacgacc ccatgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200
gaggaggcgg gggagcgggc cgccttttcc gagaggctct tcgccaacct gcttaagagg 1260
cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct 1320
gtcctggccc atatggaggc cacgggggtg cgctggacg tggcctatct cagggccttg 1380
tccctggagg tggccgagga gatcgcccg ctcgaggccg aggtcttccg cctggccggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
cttcccgcca tcggcaagac gcaaaagacc ggcaagcgct ccaccagcgc cgccgtcctg 1560
gaggccctcc gcgaggccca ccccatcggt gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgacccttg ccggacctca tccacccag gacgggcccgc 1680
ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgcgc ggccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct caggggtgctg 1860
gcccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac 1920
acggagaccg ccagctggat gttcggcgtc ccccgaggag ccgtggaccc cctgatgcgc 1980
cgggcggcca agaccatcaa cttcggggtc ctctacggca tgtcggccca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcggggggtac 2160
gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
ctcatgaagc tggtatggg gaagctcttc ccaggctgg aggaaatggg ggccaggatg 2340
ctccttcagg tcgccaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggagggtg 2460
gaggtgggga taggggagga ctggctctcc gccaggagc accaccacca ccaccac 2517

```

<210> 106
 <211> 2517
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 106
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cggggccgggc cccacgccg gaggactttc cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccag 480
gggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg 720
aagctctcct gggacctggc caaggtgcbc accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
ccccgcggg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtggggc 960
gatctttctg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg 1080
gccctgaggg aaggccttgg cctcccggcc ggcgacgacc ccatgctcct cgctacctc 1140
ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200
gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttaagagg 1260
cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct 1320
gtcctggccc atatggaggc cacgggggtg cgcctggacg tggcctatct cagggccttg 1380
tccctggagg tggccgagga gatcgccgcg ctcgaggccg aggtcttccg cctggccggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
cttcccggca tcggcaagac gcaaaagacc ggcaagcgt ccaccagcgc cgccgtcctg 1560
gaggccctcc gcgaggccca ccccatcgty gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgaccttgg ccggacctca tccaccccag gacgggcccg 1680

ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct caggggtgctg 1860
gccacacctt ccggcgacga gaacctgac cggtctctcc aggaggggag ggacatcgcc 1920
acggagaccg ccagctggat gttcggcgtc ccccgaggag ccgtggaccc cctgatgcgc 1980
cgggcgggca agaccatcaa cttcggggtc ctctacggca tgcgggcca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcggggggtac 2160
gtggagaccc tcttcggccg ccgcgcctac gtgccagacc tagaggcccg ggtgaagagc 2220
gtgcgggagg cgcccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
ctcatgaagc tggctatggt gaagctcttc ccagggtgag aggaaatggg ggccaggatg 2340
ctccttcagg tcgccaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
gcccggtgag ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
gaggtgggga taggggagga ctggctctcc gccaggagc accaccacca ccaccac 2517

<210> 107
<211> 2517
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 107
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggctctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
ccggtgcagg cggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggccgggc cccacgcgcg gaggactttc cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggctggcg cgctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc 600

ctcatgaagc tggctatggt gaagctcttc cccaggctgg aggaaatggg ggccaggatg 2340
ctccttcagg tcgccaacga gctggtcctc gagggcccaa aagagagggc ggaggccgtg 2400
gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 108
<211> 2517
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 108
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
ccggtgcagg cggctctacg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggcccggc cccacgcgcg gaggactttc cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggctggcg cgctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatactc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccggc ctggccttgg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg 720
aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
ccccgcggg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg 1080
gccctgaggg aaggccttgg cctcccgcgc ggcgacgacc ccatgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200

gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttaagagg 1260
 cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct 1320
 gtcctggccc atatggaggc cacgggggtg cgcctggacg tggcctatct cagggccttg 1380
 tccctggagg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc 1440
 cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
 cttcccgcca tcaagaagac gcaaaagacc ggcaagcgct ccaccagcgc cgccgtcctg 1560
 gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
 aagctgaaga gcacctacat tgacccttg ccggacctca tccacccag gacgggcccgc 1680
 ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
 aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggcccttcac 1800
 gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct caggggtgctg 1860
 gccacctct ccggcgacga gaacctgatc egggtcttcc aggaggggag ggacatccac 1920
 acggagaccg ccagctggat gttcggcgtc ccccgaggag ccgtggaccc cctgatgcgc 1980
 cggcgggcca agaccatcaa cttcggggtc ctctacggca tgcgggcca ccgcctctcc 2040
 caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
 gtggagaccc tcttcggccg ccgcccgtac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
 ctcatgaagc tggctatggt gaagctcttc cccaggctgg aggaaatggg ggccaggatg 2340
 ctcttcagg tcgccaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
 gcccggtg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
 gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 109
 <211> 2517
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 109
 atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggctctcct ggtggacggc 60

| | |
|--|------|
| caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag cgggggggag | 120 |
| ccggtgcagg cggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg | 180 |
| gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacgag | 240 |
| gggtacaagg cgggccgggc cccacgccg gaggactttc cccggcaact cgcctcatc | 300 |
| aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac | 360 |
| gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatcctc | 420 |
| accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag | 480 |
| gggtacctca tcaccccggc ctggccttgg gaaaagtacg gcctgaggcc cgaccagtgg | 540 |
| gccgactacc gggccctgac cggggacgag tccgacaacc tcccgggggt caagggcatc | 600 |
| ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag | 660 |
| aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccatat ggacgatctg | 720 |
| aagctctcct gggacctggc caaggtgctc accgacctgc ccctggagggt ggacttcgcc | 780 |
| aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc | 840 |
| agcctcctcc acgagttcgg ccttctggaa agcccaagg ccctggaggga ggccccctgg | 900 |
| ccccgcgcyg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcy | 960 |
| gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat | 1020 |
| aaagccctca gggacctgaa ggaggcgcyg gggcttctcg ccaaagacct gagcgttctg | 1080 |
| gccctgaggg aaggccttgg cctccgcgcc ggcgacgacc ccatgctcct cgcctacctc | 1140 |
| ctggaccctt cgaacaccac ccccagggg gtggcccggc gctacggcgg ggagtggacg | 1200 |
| gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttaagagg | 1260 |
| cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct | 1320 |
| gtcctggccc atatggaggc cacgggggtg cgcctggacg tggcctatct cagggccttg | 1380 |
| tccctggagg tggccgaggga gatcgccgc ctcgaggccg aggtcttccg cctggccggc | 1440 |
| caccccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg | 1500 |
| cttcccgcca tcggcaagac gcaaaagacc ggcaagcgct ccaccagcgc cgccgtcctg | 1560 |
| gaggccctcc gcgaggcca ccccatcgct gagaagatcc tgcagtaccg ggagctcacc | 1620 |
| aagctgaaga gcacctacat tgacccttg ccggacctca tccaccccg gacgggcccgc | 1680 |
| ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc | 1740 |
| aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgcgc ggcccttcac | 1800 |

gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct caggggtgctg 1860
 gccacctct cgggcgacga gaacctgac cgggtcttcc aggaggggag ggacatccac 1920
 acggagaccg ccagctggat gtccggcgtc ccccgaggag ccgtggaccc cctgatgcgc 1980
 cgggcggcca agaccatcaa ctccggggtc ctctacggca tgcggcccca ccgcctctcc 2040
 caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
 gtggagaccc tcttcggccg ccgcgcctac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgcgcgcgac 2280
 ctcatgaagc tggctatggt gaagctcttc cccaggctgg aggaaatggg ggccaggatg 2340
 ctcttcagg tcgccaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
 gcccggtggt ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
 gaggtgggga taggggagga ctggtctctc gccaaaggag accaccacca ccaccac 2517

<210> 110
 <211> 2517
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 110
 atgaattcgg gtagtctgcc cctctttgag cccaagggcc ggttcctcct ggtggacggc 60
 caccacctgg cctaccgcac ctccacgcc ctgaagggcc tcaccaccag cgggggggag 120
 ccggtgcagg cgtctacgg ctccgccaag agcctcctca aggcctcaa ggaggacggg 180
 gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
 gggtagaagg cgggcggggc cgagacggag gaggactttc cccggcaact cgcctcatc 300
 aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
 gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatcctc 420
 accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
 gggtagctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
 gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc 600
 ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660

| | |
|---|------|
| aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg | 720 |
| aagctctcct gggacctggc caaggtgcmc accgacctgc ccctggaggt ggacttcgcc | 780 |
| aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc | 840 |
| agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg | 900 |
| cccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtggggc | 960 |
| gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat | 1020 |
| aaagccctca gggacctgaa ggaggcgcgg gggtctctcg ccaaagacct gagcgttctg | 1080 |
| gccctgaggg aaggccttgg cctccccccc ggcgacgacc ccatgctcct cgcctacctc | 1140 |
| ctggaccctt cgaacaccac ccccgagggg gtggcccgcc gctacggcgg ggagtggacg | 1200 |
| gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttaagagg | 1260 |
| cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct | 1320 |
| gtcctggccc atatggaggc cacgggggtg cgcctggacg tggcctatct cagggccttg | 1380 |
| tccctggagg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc | 1440 |
| caccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg | 1500 |
| cttcccgcca tcggcaagac gcaaaagacc ggcaagcgt ccaccagcgc cgccgtcctg | 1560 |
| gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc | 1620 |
| aagctgaaga gcacctacat tgacctcttg ccggacctca tccaccccag gacgggccgc | 1680 |
| ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc | 1740 |
| aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggccttcac | 1800 |
| gccgaggagg ggtggtatt ggtggccctg gactatagcc agatagagct cagggtgctg | 1860 |
| gcccacctct ccggcgacga gaacctgatc cgggtcttcc aggagggggc ggacatccac | 1920 |
| acggagaccg ccagctggat gttcggcgtc cccgggagg ccgtggacc cctgatgcgc | 1980 |
| cgggcggcca agaccatcaa cttcggggtc ctctacggca tgtcggccca ccgcctctcc | 2040 |
| caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc | 2100 |
| ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac | 2160 |
| gtggagacc ttttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc | 2220 |
| gtgcgggagg cggccgagcg catggccttc aacatgccc tccagggcac cgccgccgac | 2280 |
| ctcatgaagc tggctatggt gaagctcttc ccaggtctgg aggaaatggg ggccaggatg | 2340 |
| ctccttcagg tcgccaacga gctggtcctc gaggccccaa aagagagggc ggaggccgtg | 2400 |

gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 111
<211> 2517
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 111
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
ccggtgcagg cggctctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggccctacggg 240
gggtacaagg cgggccgggc cccacgccc gaggactttc cccggcaact cgcctcatc 300
aaggagctgg tggacctcct ggggttcacg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccggc ctggctttgg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg 720
aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agccccgacc ggagaggctt agggcctttc tggagaggct tgagtttggc 840
agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
cccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accggggccc cgagccttat 1020
aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg 1080
gccctgaggg aaggccttgg cctcccgcgc ggcgacgacc ccatgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200
gaggaggcgg gggagcgggc cgccttttcc gagaggctct tcgccaacct gcttaagagg 1260

cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct 1320
gtcctggccc atatggaggc cacgggggtg cgcttgagc tggcctatct cagggccttg 1380
tccctggagg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
cttcccgcca tcggcaagac gcaaaagacc ggcaagcgt ccaccagcgc cgccgtcctg 1560
gagggcctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgaccccttg ccggacctca tccaccccag gacgggcccgc 1680
ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgcgc ggccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct caggggtgctg 1860
gcccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac 1920
acggagaccg ccagctggat gtccggcgtc ccccgaggagg ccgtggaccc cctgatgcgc 1980
cgggcgccca agaccatcaa cttcggggtc ctctacggca tgcggccca ccgctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
gtggagacc tcttcggccg ccgcgctac gtgccagacc tagaggcccg ggtgaagagc 2220
gtgcgggagg cgccgagcg catggccttc aacatgccc tccagggcac cgccgccgac 2280
ctcatgaagc tggctatggt gaagctcttc cccaggctgg aggaaatggg ggccaggatg 2340
ctccttcagg tcgccaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
gcccggtgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
gaggtgggga taggggagga ctggctctcc gccaggagc accaccacca ccaccac 2517

<210> 112
<211> 2517
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 112
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggctctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
ccggtgcagg cggtctacgg cttcgccaag agcctctca aggcctcaa ggaggacggg 180

| | |
|--|------|
| gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg | 240 |
| gggtacaagg cgggcccggc cccacgcgc gaggactttc cccggcaact cgcctcatc | 300 |
| aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac | 360 |
| gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc | 420 |
| accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag | 480 |
| gggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg | 540 |
| gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggt caagggcatc | 600 |
| ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag | 660 |
| aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccatat ggacgatctg | 720 |
| aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggagggt ggacttcgcc | 780 |
| aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc | 840 |
| agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggcctcctgg | 900 |
| ccccgcggg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc | 960 |
| gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat | 1020 |
| aaagccctca gggacctgaa ggaggcgcg gggcttctcg ccaaagacct gagcgttctg | 1080 |
| gccctgaggg aaggccttgg cctcccgcc ggcgacgacc ccatgctcct cgcctacctc | 1140 |
| ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg | 1200 |
| gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttaagagg | 1260 |
| cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct | 1320 |
| gtcctggccc atatggaggc cacgggggtg cgcctggacg tggcctatct cagggccttg | 1380 |
| tccctggagg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc | 1440 |
| caccccttca acctcaactc cggggaccag ctggaaaggg tcctctttga cgagctcagg | 1500 |
| cttcccaagt tgaagaagac gaagaagacc ggtaagcgct ccaccagcgc cgccgtcctg | 1560 |
| gaggccctcc gcgaggccca ccccatcgct gagaagatcc tgcagtaccg ggagctcacc | 1620 |
| aagctgaaga gcacctacat tgacctcttg ccggacctca tccaccccag gacgggcccgc | 1680 |
| ctccacacc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc | 1740 |
| aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggccttcac | 1800 |
| gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg | 1860 |

gcccacctct cggcgacga gaacctgatc cgggtcttcc aggaggggag ggacatccac 1920
acggagaccg ccagctggat gttcggcgtc ccccgaggag ccgtggaccc cctgatgcgc 1980
cgggcggcca agaccatcaa cttcggggtc ctctacggca tgcggccca cgcctctctc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgcgcgcgac 2280
ctcatgaagc tggctatggt gaagctcttc cccaggctgg aggaaatggg ggccaggatg 2340
ctccttcagg tcgccaacga gctggctctc gaggcccca aagagagggc ggaggccgtg 2400
gcccggtcgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 113
<211> 2517
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 113
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac ctccacgcc ctgaagggcc tcaccaccag ccggggggag 120
cgggtgcagg cggctctacg ctccgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggcccggc cccacgcgc gaggactttc cccggcaact cgcctcatc 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt cgcctcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccgcc ctggcctttg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttccgggggt caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg 720
aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780

gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 114
<211> 2517
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 114
atgaattcgg ggatgctgcc cctctttgag cccaagggcc ggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggcccggc cccacgcgcg gaggactttc cccggcaact cgcctcatc 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccgcc ctggctttgg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg 720
aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttgcc 840
agcctcctcc acgagttcgg cttcttgaa agccccaagg ccctggagga ggccccctgg 900
ccccgcggg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg 1080
gccctgaggg aaggccttgg cctcccgccc ggcgacgacc ccatgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200
gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttaagagg 1260
cttgaggggg aggagaggct ctttggtt taccgggagg tggagaggcc cttttccgct 1320
gtcctggccc atatggaggc cacgggggtg cgcctggacg tggcctatct cagggccttg 1380

cgggcggcca agaccatcaa cttcgggggtc ctctacggca tgtcggccca ccgcctctcc 2040
 caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
 gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
 ctcatgaagc tggctatggt gaagctcttc ccaggctgg aggaaatggg ggccaggatg 2340
 ctcttcagg tcgccaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
 gcccggtgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
 gagggtggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 116
 <211> 2517
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 116
 atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
 caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
 ccggtgcagg cggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
 gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
 ggttacaagg cgggccgggc cccacgcgcg gaggactttc cccggcaact cgccctcatc 300
 aaggagctgg tggacctcct ggggctggcg cgctcgagg tcccgggcta cgaggcggac 360
 gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420
 accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
 ggttacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
 gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc 600
 ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc ctcctcaag 660
 aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg 720
 aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780
 aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840

| | |
|--|------|
| agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg | 900 |
| cccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtggggc | 960 |
| gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat | 1020 |
| aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg | 1080 |
| gccctgaggg aaggccttgg cctcccgccc ggcgacgacc ccatgctcct cgcctacctc | 1140 |
| ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg | 1200 |
| gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttaagagg | 1260 |
| cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cttttccgct | 1320 |
| gtcctggccc atatggaggc cacgggggtg cgcttgacg tggcctatct cagggccttg | 1380 |
| tccctggagg tggccgagga gatcgccgc ctcgaggcgg aggtcttccg cctggccggc | 1440 |
| caccccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctcagg | 1500 |
| cttccaagt tgaagaagac gaagaagacc ggtaagcgct ccaccagcgc cgccctcctg | 1560 |
| gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc | 1620 |
| aagctgaaga gcacctacat tgacccttg ccggacctca tccaccccag gacgggccgc | 1680 |
| ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc | 1740 |
| aacctccaga acatccccgt ccgcaccccc cttgggcaga ggatccgccc ggccttcac | 1800 |
| gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct caggggtgctg | 1860 |
| gccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac | 1920 |
| acggagaccg ccagctggat gttcggcgtc ccccgaggagg ccgtggaccc cctgatgcgc | 1980 |
| cgggcggcca agaccatcaa cttcggggtc ctctacggca tgcggccca ccgcctctcc | 2040 |
| caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc | 2100 |
| ttccccaagg tgcgggcttg gattgagaag accctggagg agggcaggag gcgggggtac | 2160 |
| gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc | 2220 |
| gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac | 2280 |
| ctcatgaagc tggctatggt gaagctcttc ccaggctgg aggaaatggg ggccaggatg | 2340 |
| ctccttcagg tcgccaacga gctggtcctc gaggcccaa aagagagggc ggaggccgtg | 2400 |
| gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg | 2460 |
| gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac | 2517 |

<210> 117
 <211> 2526
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 117
 atgaattccg aggcgatgct tccgctcttt gaaccctaaag gccgggtcct cctggtggac 60
 ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
 gaaccggtgc aggcggtcta cggtctcgcc aagagcctcc tcaaggccct gaaggaggac 180
 ggggtacaagg ccgtcttcgt ggtctttgac gccaaagccc cctccttcg ccacgaggcc 240
 tacgaggcct acaaggcggg gagggccccg acccccgagg acttccccg gcagctcgcc 300
 ctcacaaagg agctggtgga cctcctgggg tttaccgccc tcgaggctcc cggctacgag 360
 gcgagacgac ttctcgccac cctggccaag aaggcggaaa aggaggggta cgaggtgagc 420
 atcctcaccg ccgaccgga cctctaccaa ctgctctccg accgctcgcc cgtcctccac 480
 cccgagggcc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag 540
 cagtgggtgg acttccgccc cctcgtgggg gacccctccg acaacctccg aggggtcagg 600
 ggcacgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc 660
 ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg 720
 gaagacctca ggctctcctt ggagctctcc cgggtgagca ccgacctccc cctggagggtg 780
 gacctcgccc aggggaggga gcccagccg gaggggctta gggccttcct ggagaggctg 840
 gaggctggca gcctcctcca cgagttcgcc ctctggagg ccccgcccc cctggaggag 900
 gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctccc ccccgagccc 960
 atgtgggagg agcttaaagc cctggccgcc tgcaggggag gccgctgca ccgggcagca 1020
 gaccccttgg cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc 1080
 gccgtcttgg cctcgaggga ggggctagac ctctgcccc gggacgaccc catgctcctc 1140
 gcctacctcc tggaccttc gaacaccacc cccgaggggg tggcgcgccg ctacgggggg 1200
 gaggggagcg agggagccgc ccaccgggccc ctctctcgg agaggctcca tcggaacctc 1260
 cttaagcgcc tcgaggggga ggagaagctc ctttggctct accacgaggt ggaaaagccc 1320
 ctctcccggg tcctggccca tatggaggcc accggggtag ggcgggacgt ggcctacctt 1380
 caggcccttt ccctggagct tgcggaggag atccgccgcc tcgaggagga ggtcttccgc 1440

| | |
|--|------|
| ttggcggggcc accccttcaa cctcaactcc cgggaccagc tggaaagggg gctctttgac | 1500 |
| gagcttaggc ttcccgctt gggaagacg caaaagacag gcaagcgctc caccagcgcc | 1560 |
| gcggtgctgg aggcctacg ggaggccac cccatcgtag agaagatcct ccagcaccgg | 1620 |
| gagctcacca agctcaagaa cacctacgtg gacccctcc caagcctcgt ccacccgagg | 1680 |
| acgggcccgc tccacacccg cttcaaccag acggccacgg ccacggggag gcttagtagc | 1740 |
| tccgacccca acctgcagaa catccccgtc cgcacccctc tgggccagag gatccgccgg | 1800 |
| gccttcgtgg ccgaggcggg ttgggcgttg gtggcccttg actatagcca gatagagctc | 1860 |
| cgcgtcctcg cccacctctc cggggacgaa aacctgatca gggctctcca ggaggggaag | 1920 |
| gacatccaca cccagaccgc aagctggatg ttcggcgtcc ccccgagggc cgtggacccc | 1980 |
| ctgatgcgcc gggcgcccaa gacggtgaac ttcggcgtcc tctacggcat gtccgcccat | 2040 |
| aggctctccc aggagcttgc catcccctac gaggaggcgg tggcctttat agagcgctac | 2100 |
| ttccaaagct tccccaaggt gcgggccttg atagaaaaga ccctggagga ggggaggaag | 2160 |
| cggggctacg tggaaaccct cttcggaaga aggcgctacg tgcccgacct caacgcccgg | 2220 |
| gtgaagagcg tcagggaggc cgcggagcgc atggccttca acatgcccgt ccagggcacc | 2280 |
| gccgccgacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg | 2340 |
| gcccgcatgc tcctccaggt cgccaacgag ctctccttg agggccccca agcgcggggc | 2400 |
| gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatccctc cgccgtgccc | 2460 |
| ctggaggtgg aggtggggat gggggaggac tggctttccg ccaagggcca ccaccaccac | 2520 |
| caccac | 2526 |

<210> 118
 <211> 2526
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

| | |
|---|-----|
| <400> 118 | |
| atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac | 60 |
| ggccaccacc tggcctaccg caccttcttc gcctgaagg gcctcaccac gagccggggc | 120 |
| gaaccggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac | 180 |
| gggtacaagg ccgtcttcgt ggtctttgac gccaaaggccc cctccttcctg ccacgaggcc | 240 |
| tacgaggcct acaaggcggg gagggccccg acccccgagg acttcccccg gcagctcgcc | 300 |

| | |
|---|------|
| ctcatcaagg agctggtgga cctcctggggg tttaccgcgc tcgagggtccc cggctacgag | 360 |
| gcggacgacg ttctcgccac cctggccaag aaggcggaag aggaggggta cgagggtgcgc | 420 |
| atcctcaccg ccgaccgcga cctctaccaa ctctgtctccg accgcgtcgc cgtcctccac | 480 |
| cccgagggcc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag | 540 |
| cagtgggtgg acttccgcgc cctcgtggggg gacccctccg acaacctccc cgggggtcaag | 600 |
| ggcatcgggg agtataccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc | 660 |
| ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg | 720 |
| gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctccc cctggagggtg | 780 |
| gacctcgcgc aggggcggga gcccgcgcgc gaggggctta gggccttctt ggagagggtg | 840 |
| gagttcggca gcctcctcca cgagttcggc ctcttgaggg ccccgcccc cctggaggag | 900 |
| gccccctggc ccccgccgga aggggccttc gtgggcttcg tctctctccg ccccgagccc | 960 |
| atgtgggagg agcttaaagc cctggccgcc tgcaggggag gccgcgtgca ccgggcagca | 1020 |
| gaccccttgg cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc | 1080 |
| gccgtcttgg cctcgaggga ggggctagac ctctgcccc gggacgaccc catgctcttc | 1140 |
| gcctacctcc tggaccttc gaacaccacc cccgaggggg tggcgcgccg ctacgggggg | 1200 |
| gagtggacgg aggacgccgc ccaccgggccc ctctctcgg agaggctcca tcggaacctc | 1260 |
| cttaagcgcc tcgaggggga ggagaagctc ctttggtctt accacgaggt ggaaaagccc | 1320 |
| ctctcccggg tcttgccca tatggaggcc accggggtac ggcgggacgt ggcctacctt | 1380 |
| caggcccttt ccctggagct tgcggaggag atccgccgcc tcgaggagga ggtcttccgc | 1440 |
| ttggcgggccc accccttcaa cctcaactcc cgggaccagc tggaaagggt gctctttgac | 1500 |
| gagcttaggc ttcccgcctt ggggaagacg caaaagacag gcaagcgctc caccagcgcc | 1560 |
| gcggtgctgg aggccctacg ggaggccac cccatcgtgg agaagatcct ccagcaccgg | 1620 |
| gagctacca agctcaagaa cacctacgtg gacccctcc caagcctcgt ccacccgagg | 1680 |
| acgggcccgc tccacacccg cttcaaccag acggccacgg ccacggggag gcttagtagc | 1740 |
| tccgacccca acctgcagaa catccccgtc cgcacccctt tgggcccagag gatccgcccg | 1800 |
| gccttcgtgg ccgaggcggg ttgggcgttg gtggccctgg actatagcca gatagagctc | 1860 |
| cgctcctcg cccacctctc cggggacgaa aacctgatca gggctctcca ggaggggaag | 1920 |
| gacatccaca cccagaccgc aagctggatg ttcggcgctc ccccgagggc cgtggacccc | 1980 |

| | |
|---|------|
| gagttcggca gcctcctcca cgagttcggc ctcttgagg ccccgcccc cctggaggag | 900 |
| gccccctggc ccccgccgga aggggccttc gtgggcttcg tctctctccg ccccgagccc | 960 |
| atgtggggcg agcttaaagc cctggccgcc tgcaggggag gccgcgtgca ccgggcagca | 1020 |
| gacccttgg cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc | 1080 |
| gccgtcttgg cctcgaggga ggggctagac ctctgcccc gggacgacct catgctcctc | 1140 |
| gcctacctcc tggaccttc gaacaccacc cccgaggggg tggcgcgggc ctacgggggg | 1200 |
| gagtggacgg aggacgccgc ccaccgggcc ctctctcgg agaggctcca tcggaacctc | 1260 |
| cttaagcgcc tcgaggggga ggagaagctc ctttggtctt accacgaggt ggaaaagccc | 1320 |
| ctctccccgg tcttgcccc tatggaggcc accggggtac ggcgggacgt ggcctacctt | 1380 |
| caggcccttt ccttgagct tgcggaggag atccgccgc tcgaggagga ggtcttccgc | 1440 |
| ttggcgggcc accccttcaa cctcaactcc cgggaccagc tggaaagggg gctctttgac | 1500 |
| gagcttaggc ttccgcctt ggggaagacg caaaagacag gcaagcgctc caccagcgcc | 1560 |
| gcggtgctgg aggccttac ggaggccac cccatcgtag agaagatcct ccagcaccgg | 1620 |
| gagctacca agctcaagaa cacctacgtg gacccctcc caagcctcgt ccacccgagg | 1680 |
| acgggcccgc tccacaccgc cttcaaccag acggccacgg ccacggggag gcttagtagc | 1740 |
| tccgaccca acctgcagaa catccccgtc cgcacccctc tgggcccagag gatccgcggg | 1800 |
| gccttcgtgg ccgaggcggg ttgggcgttg gtggccctgg actatagcca gatagagctc | 1860 |
| cgcgtcctcg cccacctctc cggggacgaa aacctgatca gggctctcca ggaggggaag | 1920 |
| gacatccaca cccagaccgc aagctggatg ttcggcgctc ccccgagggc cgtggacccc | 1980 |
| ctgatgcgcc gggcgcccaa gacggtgaac ttcggcgctc tctacggcat gtccgcccac | 2040 |
| aggctctccc aggagcttgc catcccctac gaggaggcgg tggcctttat agagcgctac | 2100 |
| ttccaaagct tcccaaggt gcgggcctgg atagaaaaga ccctggagga ggggaggaag | 2160 |
| cggggctacg tggaaaccct cttcggaaga aggcgctacg tgcccacact caacgcccgg | 2220 |
| gtgaagagcg tcagggaggc cgcggagcgc atggccttca acatgcccgt ccagggcacc | 2280 |
| gccgccgacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg | 2340 |
| gcccgcatgc tctccaggt cgccaacgag ctctcctcgg agggccccca agcgcgggcc | 2400 |
| gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatcccct cgccgtgccc | 2460 |
| ctggaggtgg aggtggggat gggggaggac tggctttccg ccaagggtca ccaccaccac | 2520 |

caccac

2526

<210> 120
<211> 2517
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 120
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtcct tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cggggccgggc cccacgcgcg gaggactttc cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggttcacg cgctcagagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccggc ctggctttgg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg 720
aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
agcctcctcc acgagttcgg cttcttgaa agccccaagg ccctggagga ggccccctgg 900
cccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg 1080
gccctgaggg aaggccttgg cttcccgccc ggcgacgacc ccattgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200
gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttaagagg 1260
cttgaggggg aggagaggct cttttggctt taccgggagg tggagaggcc cttttccgct 1320
gtcctggccc atatggaggc cacgggggtg cgcttgagcg tggcctatct cagggccttg 1380

tccctggagg tggccgagga gatcgcccg ctcgaggccg aggtcttccg cctggccggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctcagg 1500
cttcccaagt tgaagaagac gaagaagacc ggtaagcgct ccagcagcgc cgccgtcctg 1560
gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgaccccttg ccggacctca tccacccag gacgggcccgc 1680
ctccacaccc gttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct caggggtgctg 1860
gcccacctct ccggcgacga gaacctgac cgggtcttcc aggaggggcg ggacatccac 1920
acggagaccg ccagctggat gttcggcgtc ccccgggagg ccgtggaccc cctgatgcgc 1980
cgggcgccca agaccatcaa cttcggggtc ctctacggca tgtcggccca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccaagg tgcgggcttg gattgagaag accctggagg agggcaggag gcgggggtac 2160
gtggagaccc tcttcggccg ccgccgtac gtgccagacc tagaggcccg ggtgaagagc 2220
gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
ctcatgaagc tggctatggt gaagctcttc ccaggctgg aggaaatggg ggccaggatg 2340
ctccttcagg tcgccaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 121
<211> 2517
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 121
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
ccggtgcagg cggctctacgg cttcgccaag agcctctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240

cgggcggcca agaccatcaa cttcggggtc ctctacggca tgcgggcca ccgcctctcc 2040
 caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcggggggtac 2160
 gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
 ctcatgaagc tggctatggt gaagctcttc cccaggctgg aggaaatggg ggccaggatg 2340
 ctcttccagg tcgccaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
 gcccggtg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
 gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 122
 <211> 2517
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 122
 atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggctctcct ggtggacggc 60
 caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
 ccggtgcagg cggtctacgg cttcgccaaag agcctcctca aggccctcaa ggaggacggg 180
 gacgcggtga tcgtggtctt tgacgccaaag gccccctcct tccgccacga ggcctacggg 240
 ggggtacaagg cgggccggggc cccacgcgcg gaggactttc cccggcaact cgccctcatc 300
 aaggagctgg tggacctcct ggggttcacg cgcctcgagg tcccgggcta cgaggcggac 360
 gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatcctc 420
 accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
 gggtagctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
 gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caatggcatc 600
 ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc ctcctcaag 660
 aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg 720
 aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780
 aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840

| | | | | | | |
|------------|------------|-------------|------------|------------|-------------|------|
| agcctcctcc | acgagttcgg | ccttctggaa | agccccaagg | ccctggagga | ggccccctgg | 900 |
| ccccgcgagg | aaggggcctt | cgtgggcttt | gtgctttccc | gcaaggagcc | catgtggggc | 960 |
| gatcttctgg | ccctggccgc | cgccaggggc | ggccgcgtcc | accggggccc | cgagccttat | 1020 |
| aaagccctca | gggacctgaa | ggaggcgagg | gggcttctcg | ccaaagacct | gagcgttctg | 1080 |
| gccctgaggg | aaggccttgg | cctcccggcc | ggcgacgacc | ccatgctcct | cgcttacctc | 1140 |
| ctggaccctt | cgaacaccac | ccccgagggg | gtggcccggc | gctacggcgg | ggagtggacg | 1200 |
| gaggaggcgg | gggagcgggc | cgccctttcc | gagaggctct | tcgccaacct | gcttaagagg | 1260 |
| cttgaggggg | aggagaggct | cctttggctt | taccgggagg | tggagaggcc | cctttccgct | 1320 |
| gtcctggccc | atatggaggc | cacgggggtg | cgctggacg | tggcctatct | cagggccttg | 1380 |
| tccctggagg | tggccgagga | gatcgcccgc | ctcgaggccg | aggtcttccg | cctggccggc | 1440 |
| caccccttca | acctcaactc | ccgggaccag | ctggaaaggg | tcctctttga | cgagctcagg | 1500 |
| cttcccaagt | tgaagaagac | gaagaagacc | ggtaagcgct | ccagcagcgc | cgccgtcctg | 1560 |
| gaggccctcc | gcgaggccca | ccccatcgct | gagaagatcc | tgcagtaccg | ggagctcacc | 1620 |
| aagctgaaga | gcacctacat | tgaccccttg | ccggacctca | tccaccccag | gacgggcccgc | 1680 |
| ctccacaccc | gcttcaacca | gacggccacg | gccacgggca | ggctaagtag | ctccgatccc | 1740 |
| aacctccaga | acatccccgt | ccgcaccccg | cttgggcaga | ggatccggcg | ggccttcctc | 1800 |
| gccgaggagg | ggtggctatt | ggtggccctg | gactatagcc | agatagagct | caggggtgctg | 1860 |
| gcccacctct | ccggcgacga | gaacctgatc | cggttcttcc | aggagggggc | ggacatccac | 1920 |
| acggagaccg | ccagctggat | gttcggcgct | ccccgggagg | ccgtggaccc | cctgatgcgc | 1980 |
| cgggcggcca | agaccatcaa | cttcggggct | ctctacggca | tgtcggccca | ccgcctctcc | 2040 |
| caggagctag | ccatccctta | cgaggaggcc | caggccttca | ttgagcgcta | ctttcagagc | 2100 |
| ttccccaagg | tgcgggcctg | gatttgagaag | accctggagg | agggcaggag | gccccgggtac | 2160 |
| gtggagaccc | tcttcggccg | ccgccgctac | gtgccagacc | tagaggcccg | ggtgaagagc | 2220 |
| gtgcgggagg | cggccgagcg | catggccttc | aacatgcccg | tccagggcac | cgccgccgac | 2280 |
| ctcatgaagc | tggctatggt | gaagctcttc | cccaggctgg | aggaaatggg | ggccaggatg | 2340 |
| ctccttcagg | tcgccaacga | gctggctctc | gaggcccaaa | aagagagggc | ggaggccgtg | 2400 |
| gcccggtcgg | ccaaggaggt | catggagggg | gtgtatcccc | tggccgtgcc | cctggaggtg | 2460 |
| gaggtgggga | taggggagga | ctggctctcc | gccaaggagc | accaccacca | ccaccac | 2517 |

<210> 123
 <211> 2517
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 123
 atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
 caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
 ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
 gacgcggtga tcgtggtcct tgacgccaag gccccctcct tccgccacga ggcctacggg 240
 ggggtacaagg cgggcccgggc cccacgcccg gaggactttc cccggcaact cgcctcatc 300
 aaggagctgg tggacctcct ggggttcacg cgctcgagg tcccgggcta cgaggcggac 360
 gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatcctc 420
 accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
 ggggtacctca tcaccccggc ctggccttgg gaaaagtacg gcctgaggcc cgaccagtgg 540
 gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggt caagggcatc 600
 ggggagaaga cgcagaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
 aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggccacat ggacgatctg 720
 aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggagggt ggacttcgcc 780
 aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtctggc 840
 agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
 ccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc 960
 gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
 aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg 1080
 gccctgaggg aaggccttgg cctcccgcgc ggcgacgacc ccatgctcct cgcctacctc 1140
 ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200
 gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttaagagg 1260
 cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct 1320
 gtcttgggcc atatggaggc cacgggggtg cgcttgagc tggcctatct cagggccttg 1380
 tccttgaggg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc 1440

caccccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctcagg 1500
cttcccaagt tgaagaagac gaagaagacc ggtaagcgct ccagcagcgc cgccgtcctg 1560
gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgaccccttg ccggacctca tccaccccag gacgggcccgc 1680
ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
gcccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggag ggacatccac 1920
acggagaccg ccagctggat gttcggcgtc ccccgaggag ccgtggaccc cctgatgcgc 1980
cgggcggcca agaccatcaa cttcggggtc ctctacggca tgtcggccca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
ctcatgaagc tggctatggt gaagctcttc ccagggtgg aggaaatggg ggccaggatg 2340
ctccttcagg tcgccaacga gctggctctc gaggcccca aagagagggc ggaggccgtg 2400
gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggag accaccacca ccaccac 2517

<210> 124
<211> 2517
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 124
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
ccggtgcagg cggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggccgggc cccacgcccg gaggactttc cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggttcacg cgctcgagg tcccgggcta cgaggcggac 360

| | |
|---|------|
| gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatcctc | 420 |
| accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag | 480 |
| gggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg | 540 |
| gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatg | 600 |
| ggggagaaga cggggaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag | 660 |
| aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccatat ggacgatctg | 720 |
| aagctctcct gggacctggc caaggtgctc accgacctgc ccctggaggt ggacttcgcc | 780 |
| aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc | 840 |
| agcctcctcc acgagttcgg ctttctggaa agccccaagg ccctggagga ggccccctgg | 900 |
| ccccgcgcgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtggggc | 960 |
| gatcttcttg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat | 1020 |
| aaagccctca gggacctgaa ggaggcgcg gggtctctcg ccaaagacct gagcgttctg | 1080 |
| gccctgaggg aaggccttgg cctccgcgcc ggcgacgacc ccatgctcct cgcctacctc | 1140 |
| ctggaccctt cgaacaccac ccccagggg gtggcccggc gctacggcgg ggagtggacg | 1200 |
| gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttaagagg | 1260 |
| cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct | 1320 |
| gtcctggccc atatggaggc cacgggggtg cgcctggacg tggcctatct cagggccttg | 1380 |
| tccttgagg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc | 1440 |
| cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctcagg | 1500 |
| cttcccaagt tgaagaagac gaagaagacc ggtaagcgct ccagcagcgc cgccgtcctg | 1560 |
| gaggccctcc gcgaggccca ccccatcggt gagaagatcc tgcagtaccg ggagctcacc | 1620 |
| aagctgaaga gcacctacat tgacccttg ccggacctca tccacccag gacgggcccgc | 1680 |
| ctccacacc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc | 1740 |
| aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccc ggccttcac | 1800 |
| gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg | 1860 |
| gccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac | 1920 |
| acggagaccg ccagctggat gttcggcgtc ccccgaggag ccgtggaccc cctgatgcgc | 1980 |
| cggggggcca agaccatcaa cttcggggtc ctctacggca tgtcggccca ccgcctctcc | 2040 |

caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
 gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgcgcgcgac 2280
 ctcatgaagc tggctatggt gaagctcttc cccaggctgg aggaaatggg ggccaggatg 2340
 ctccttcagg tcgccaacga gctggtcctc gaggcccaa aagagagggc ggaggccgtg 2400
 gcccggtg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggagggtg 2460
 gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 125
 <211> 2517
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 125
 atgaattcgg ggatgctgcc cctctttgag cccaagggcc ggtcctcct ggtggacggc 60
 caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
 ccggtgcagg cggctctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
 gacgcggtga tcgtggtctt tgacgccaaag gccccctcct tccgccacga ggcttacggg 240
 ggggtacaagg cgggcccggc cccacgcgcg gaggactttc cccggcaact cgcctcatc 300
 aaggagctgg tggacctcct ggggttcacg cgcctcgagg tcccgggcta cgaggcggac 360
 gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatcctc 420
 accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
 gggtagctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
 gccgactacc gggccctgac cggggacgag tccgacaacc tcccgggggt caagggcatc 600
 ggggagaata cggcgaggaa gcttctggag gagggggga gcctggaagc ctcctcaag 660
 aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggccacat ggacgatctg 720
 aagctctcct gggacctggc caaggcgcg accgacctgc ccctggagggt ggacttcgcc 780
 aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
 agcctcctcc acgagttcgg cttcttgaa agccccaagg ccctggagga ggccccctgg 900
 cccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc 960

gatcttcttg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
aaagccctca gggacctgaa ggaggcgcg gggtctctcg ccaaagacct gagcgttctg 1080
gccctgaggg aaggccttgg cctcccgccc ggcgacgacc ccatgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200
gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttaagagg 1260
cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct 1320
gtcctggccc atatggaggc cacgggggtg cgcttgacg tggcctatct cagggccttg 1380
tccctggagg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tctctttga cgagctcagg 1500
cttcccaagt tgaagaagac gaagaagacc ggtaagcgct ccagcagcgc cgccgtcctg 1560
gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgacccttg ccggacctca tccacccag gacgggcccgc 1680
ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcacccg cttgggcaga ggatccgccc ggcttccatc 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
gcccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac 1920
acgggagaccg ccagctggat gttcggcgtc cccggggagg ccgtggaccc cctgatgcgc 1980
cgggcggcca agaccatcaa cttcggggtc ctctacggca tgtcggcca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
ctcatgaagc tggctatggt gaagctcttc ccaggctgg aggaaatggg ggccaggatg 2340
ctccttcagg tcgccaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
gcccggtg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 126
<211> 2517
<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 126

| | |
|--|------|
| atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc | 60 |
| caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag | 120 |
| ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg | 180 |
| gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccaaga ggcctacggg | 240 |
| gggtacaagg cgggcccgggc cccacgccg gaggactttc cccggcaact cgcctcatc | 300 |
| aaggagctgg tggacctcct ggggttcacg cgcctcgagg tcccgggcta cgaggcggac | 360 |
| gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc | 420 |
| accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccag | 480 |
| gggtacctca tcaccccggc ctggccttgg gaaaagtacg gcctgaggcc cgaccagtgg | 540 |
| gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc | 600 |
| ggggagaagc cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag | 660 |
| aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg | 720 |
| aagctctcct gggacctggc caagggtgcg accgacctgc ccctggaggt ggacttcgcc | 780 |
| aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc | 840 |
| agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg | 900 |
| ccccgcggg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc | 960 |
| gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat | 1020 |
| aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg | 1080 |
| gccctgaggg aaggccttgg cctcccgcc ggcgacgacc ccatgctcct cgctacctc | 1140 |
| ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg | 1200 |
| gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttaagagg | 1260 |
| cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cttttccgct | 1320 |
| gtcctggccc atatggaggc cacgggggtg cgcttgagc tggcctatct cagggccttg | 1380 |
| tccctggagg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc | 1440 |
| caccccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctcagg | 1500 |
| cttcccaagt tgaagaagac gaagaagacc ggtaagcgct ccagcagcgc cgccgtcctg | 1560 |

gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgacccttg ccggacctca tccacccag gacgggcccgc 1680
ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggcccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct caggggtgctg 1860
gcccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggagc ggacatccac 1920
acggagaccg ccagctggat gttcggcgtc ccccgaggag ccgtggaccc cctgatgcgc 1980
cgggcggcca agaccatcaa cttcggggtc ctctacggca tgtcggccca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgcgcggac 2280
ctcatgaagc tggctatggt gaagctcttc cccaggctgg aggaaatggg ggccaggatg 2340
ctccttcagg tcgccaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 127
<211> 2517
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 127
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggctcctct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
ccggtgcagg cggcttacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggcccggc cccacgccc gaggactttc cccggcaact cgcctcatc 300
aaggagctgg tggacctcct ggggttcacg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420

| | |
|--|------|
| accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag | 480 |
| gggtacctca tcaccccggc ctggccttgg gaaaagtacg gcctgaggcc cgaccagtgg | 540 |
| gccgactacc gggccctgac cggggacgag tccgacaacc ttcccgggggt caagggcatc | 600 |
| ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag | 660 |
| aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg | 720 |
| aagctctcct gggacctggc caaggtgcg cccgacctgc ccctggagggt ggacttcgcc | 780 |
| aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagagggt tgagtttggc | 840 |
| agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg | 900 |
| cccccgccgg aaggggcctt cgtgggcctt gtgctttccc gcaaggagcc catgtgggccc | 960 |
| gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat | 1020 |
| aaagccctca gggacctgaa ggaggcgcg gggcttctcg ccaaagacct gagcgttctg | 1080 |
| gccctgaggg aaggccttgg cctcccgccc ggcgacgacc ccatgctcct cgccctacctc | 1140 |
| ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg | 1200 |
| gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttaagagg | 1260 |
| cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct | 1320 |
| gtcctggccc atatggaggc cacgggggtg cgctggacg tggcctatct cagggccttg | 1380 |
| tccctggagg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc | 1440 |
| caccccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctcagg | 1500 |
| cttcccaagt tgaagaagac gaagaagacc ggtaagcgct ccagcagcgc cgccgtcctg | 1560 |
| gaggccctcc gcgaggccca ccccatcggtg gagaagatcc tgcagtaccg ggagctcacc | 1620 |
| aagctgaaga gcacctacat tgacccttg ccggacctca tccaccccag gacgggcccgc | 1680 |
| ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc | 1740 |
| aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgcgg ggccttcac | 1800 |
| gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg | 1860 |
| gcccacctct ccggcgacga gaacctgac cggtcttccc aggagggggc ggacatccac | 1920 |
| acgggagacc ccagctggat gttcggcgtc ccccgggagg ccgtggaccc cctgatgcgc | 1980 |
| cgggcgggca agaccatcaa cttcggggtc ctctacggca tgtcggccca cgccctctcc | 2040 |
| caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc | 2100 |
| ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac | 2160 |

gtggagaccc tcttcggccg ccgccgtac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgctggagg cgcccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
 ctcatgaagc tggctatggt gaagctcttc cccaggctgg aggaaatggg ggccaggatg 2340
 ctcttcagg tcgccaacga gctggctctc gagggcccaa aagagagggc ggaggccgtg 2400
 gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggagggtg 2460
 gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 128
 <211> 2517
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 128
 atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
 caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
 ccggtgcagg cggctacagg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
 gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
 gggtagaagg cgggccgggc ccccacgccg gaggactttc cccggcaact cgccctcatc 300
 aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
 gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatcctc 420
 accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccag 480
 gggtagctca tcaccccgcc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
 gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcac 600
 ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
 aacctggacc ggctgaagcc cgccatcccg gagaagatcc tggcccacat ggacgatctg 720
 aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggagggt ggacttcgcc 780
 aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttgcc 840
 agcctcctcc acgagttcgg ctttctggaa agccccaagg ccctggagga gggccctgg 900
 ccccgcggc aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc 960
 gatcttctgg ccctggccgc cgccaggggc ggccgcgtgc accgggcagc agacccttg 1020

gcgggggctaa aggacctcaa ggaggtccgg ggctcctcgc ccaaggacct cgccgtcttg 1080
gcctcgaggg aggggctaga cctcgtgccc ggggacgacc ccatgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccgagggg gtggcgcggc gctacggggg ggagtggacg 1200
gaggacgccc cccaccgggc cctcctctcg gagaggctcc atcggaacct ccttaagcgc 1260
ctcgaggggg aggagaagct cctttggctc taccacgagg tggaaaagcc cctctcccgg 1320
gtcctggccc atatggaggc caccggggta cggcgggacg tggcctacct tcaggccctt 1380
tccctggagc ttgcggagga gatccgccgc ctcgaggagg aggtcttccg cttggcgggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tgctctttga cgagcttagg 1500
cttcccgcct tgaagaagac gaagaagaca ggcaagcgct ccaccagcgc cgcggtgctg 1560
gaggccctac gggaggccca ccccatcgtg gagaagatcc tccagcaccg ggagctcacc 1620
aagctcaaga acacctacgt ggacccccctc ccaagcctcg tccacccgag gacgggcccgc 1680
ctccacaccc gcttcaacca gacggccacg gccacgggga ggcttagtag ctccgacccc 1740
aacctgcaga acatccccgt ccgcaccccc ttggggccaga ggatccgccg ggccttcgtg 1800
gccgaggcgg gttgggcgtt ggtggccctg gactatagcc agatagagct ccgcgtcctc 1860
gccacctct ccggggacga aaacctgatc agggctcttc aggaggggaa ggacatccac 1920
accagaccg caagctggat gttcggcgtc ccccgagg cctgggacct cctgatgcgc 1980
cgggcggcca agacggtgaa cttcggcgtc ctctacggca tgtccgcca taggctctcc 2040
caggagcttg ccatccccta cgaggaggcg gtggccttta tagagcgcta cttccaaagc 2100
ttcccaagg tgcgggcctg gatagaaaag accctggagg aggggaggaa gcggggctac 2160
gtggaaacct tcttcggaag aaggcgctac gtgcccgaac tcaacgcccg ggtgaagagc 2220
gtcagggagg ccgcggagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
ctcatgaagc tcgccatggt gaagctcttc cccgcctcc gggagatggg ggcccgcatg 2340
ctcctccagg tcgccaacga gctcctcctg gaggcccccc aagcgcgggc cgaggagggtg 2400
gcggcttttg ccaaggaggc catggagaag gcctatcccc tcgccgtgcc cctggagggtg 2460
gaggtgggga tgggggagga ctggctttcc gccaaagggtc accaccacca ccaccac 2517

<210> 129
<211> 2526
<212> DNA
<213> Artificial
<220>

<223> Synthetic

<400> 129

| | |
|--|------|
| atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac | 60 |
| ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc | 120 |
| gaaccggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac | 180 |
| gggtacaagg ccgtcttcgt ggtctttgac gccaaaggccc cctccttcg ccacgaggcc | 240 |
| tacgaggcct acaaggcggg gagggccccg acccccgagg acttcccccg gcagctcgcc | 300 |
| ctcatcaagg agctggtgga cctcctgggg tttaaccgcc tcgagggtccc cggctacgag | 360 |
| gcggacgacg ttctcgccac cctggccaag aaggcggaaa aggaggggta cgagggtgcg | 420 |
| atcctcaccg ccgaccgcga cctctaccaa ctctgtctcg accggtcgc cgtcctccac | 480 |
| cccgagggcc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag | 540 |
| cagtgggtgg acttccgcgc cctcgtgggg gacctctcg acaacctccc cgggggtcaag | 600 |
| ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc | 660 |
| ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg | 720 |
| gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctccc cctggagggtg | 780 |
| gacctcgccc aggggcggga gcccgaccgg gaggggctta gggccttcct ggagaggctg | 840 |
| gagttcggca gcctcctcca cgagttcggc ctcttgagg ccccgcccc cctggaggag | 900 |
| gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctcccg ccccgagccc | 960 |
| atgtgggcgg agcttaaagc cctggccgcc tgcagggcg gccgcgtcca ccgggcccc | 1020 |
| gagccttata aagccctcag ggacctgaag gaggcgcggg ggcttctcgc caaagacctg | 1080 |
| agcgttctgg ccctgaggga aggccttggc ctccgcccc gcgacgacct catgctcctc | 1140 |
| gcctacctcc tggaccttc gaacaccacc cccgaggggg tggcccgcg ctacggcggg | 1200 |
| gagtggacgg aggaggcggg ggagcgggccc gccctttccg agaggctctt cgccaacctg | 1260 |
| cttaagaggc ttgaggggga ggagaggctc ctttggcttt accgggagggt ggagaggccc | 1320 |
| ctttccgctg tcctggccca tatggaggcc acgggggtgc gcctggacgt ggcctatctc | 1380 |
| agggccttgt ccctggagggt ggccgaggag atcgcccgcc tcgaggccga ggtcttccgc | 1440 |
| ctggccggcc accccttcaa cctcaactcc cgggaccagc tggaaagggt cctctttgac | 1500 |
| gagctagggc ttcccgccat caagaagacg caaaagaccg gcaagcgctc caccagcgcc | 1560 |
| gccgtcctgg aggcctccg cgaggcccac cccatcgtgg agaagatcct gcagtaccgg | 1620 |

gggaccctgg tcaccccaaa ggacgtccag gagaagtacg ggggtgcccc ggagcgctgg 540
 gtggacttcc gcgccctcac gggggaccgc tcggacaaca tccccggggt ggcggggata 600
 ggggagaaga ccgcccttcg actcctcgca gagtggggga gcgtggaaaa cctcctgaag 660
 aacctggacc gggtaaagcc ggactcgctc cggcgcaaga tagaggcgca cctcgaggac 720
 ctccacctct ccttagacct ggcccgcac cgcaccgacc tccccctgga ggtggacttt 780
 aaggccctgc gccgcaggac ccccgacctg gagggcctga gggccttttt ggaggagctg 840
 gagttcggaa gcctcctcca cgagttcggc ctctctggag gggagaagcc ccgggaggag 900
 gccccctggc ccccgccga aggggccttc gtgggcttcc tcctttcccg caaggagccc 960
 atgtgggcgg agcttctggc cctggcgccg gcctcggagg gccgggtcca ccgggcaaca 1020
 agcccggttg aggccttggc cgacctcaag gagggccggg ggttcctggc caaggacctg 1080
 gccgttttgg cctgcggga ggggggtggc ctggacccca cggacgacct cctcctggtg 1140
 gcctacctcc tggaccggc caacacccac cccgaggggg tggcccgcg ctacgggggc 1200
 gagttcacgg aggacgcagc ggagagggcc ctctctccg agaggctctt ccagaacctc 1260
 tttaaacggc tttccgagaa gctcctctgg ctctaccagg aggtggagcg gccctctcc 1320
 cgggtcttgg ccacatgga gggccggggg gtgaggctgg acgtccccct tctggaggcc 1380
 ctctcctttg agctggagaa ggagatggag cgctcggagg gggaggctct ccgtttggcc 1440
 ggccaccct tcaacctcaa ctcccgac cagctggaaa gggctctctt tgacgagctg 1500
 ggcctcacc cgggtgggccc gacgcagaag acgggcaagc gctccaccgc ccagggggcc 1560
 ctggaggccc tccggggggc ccacccatc gtggagctca tcctccagta ccgggagctt 1620
 tccaagctca aaagcaccta cctggacccc ctgcccggc tcgtccacc gcggacgggc 1680
 cggctccaca ccgcttcaa ccagacggcc acggccacgg gaaggctttc cagctccgac 1740
 cccaacctgc agaacatccc cgtgcgcacc cccttggggc agcgcatccg caaggccttc 1800
 gtggccgagg aggggtggct ccttttggcg gcggactact ccagattga gctccgggtc 1860
 ctggcccacc tctcggggga cgagaacctg aagcgggtct tccgggagg gaaggacatc 1920
 cataccgaga ccgcccctg gatgttcggc ttagaccccg ctctggtgga tccaaagatg 1980
 cgccggcgcg ccaagacggt caacttcggc gtcctctacg ggatgtccgc ccacaggctc 2040
 tcccaggagc tcggcataga ctacaaggag gcggaggcct ttattgagcg ctacttccag 2100
 agcttcccca aggtgcgggc ctggatagaa aggaccctgg aggagggccg gacgcggggc 2160

tacgtggaga ccctgttcgg caggaggcgc tatgtgcccg acctggcctc ccgggtccgc 2220
 tcggtgcggg aggcggcgga gcggatggcc ttcaacatgc ccgtgcaggg caccgccgcc 2280
 gacctgatga agatcgccat ggtcaagctc ttccccaggg taaagcccct gggggcccac 2340
 ctctcctcc aagtgcacaa cgagctggtc ctggaggtgc ccgaggaccg ggccgaggag 2400
 gccaaaggccc tggtaagga ggtcatggag aacgcctacc ccctggacgt gcccctcgag 2460
 gtggaggtgg gcgtgggtcg ggactggctg gaggcgaagc aggattga 2508

<210> 131
 <211> 2499
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 131
 atggaattcc tgcccctctt tgagcccaag ggccgggtgc ttctgggtga cggccaccac 60
 ctggcctacc gtaccttttt tgccctgaag ggcctcacca ccagccgcgg ggagccggtc 120
 caggcgggtgt acgggtttgc caagagcctt ttgaaggcgc taagggaaga cggggatgtg 180
 gtgatcgtgg tgtttgacgc caaggccccc tccttcgcgc accagaccta cgaggcctac 240
 aaggcggggc gggctccac ccccaggac tttcccggc agcttgccct tatcaaggag 300
 atggtggacc ttttgggctt ggagcgctc gaggtgccgg gctttgaagc ggatgacgtc 360
 ctggctaccc tggccaagaa ggcgaaaag gaaggctacg aagtgcgcat cctcaccgcg 420
 gaccgggacc tttaccagct tctttcggag cgaatctcca tccttcaccc ggagggttac 480
 ctgatcaccc cggagtggct ttgggagaag tatgggctta agccttccca gtgggtggac 540
 taccgggcct tggccgggga cccttcgcg aacatccccg gcgtgaaggg catcggggag 600
 aagacggcgg ccaagctgat ccgggagtgg ggaagcctgg aaaaccttct taagcacctg 660
 gaacaggtga aacctgcctc cgtgcgggag aagatcctta gccacatgga ggacctcaag 720
 ctatccctgg agctatcccg ggtgcacacg gacttgctcc ttcaggtgga cttcgcccgg 780
 cgccgggagc cggaccggga ggggcttaag gccttttttg agaggctgga gttcggaagc 840
 ctctccacg agttcggcct gttggaaagc ccggtggcgg cggaggaagc tccttgcccg 900
 cccccgagg gagccttcgt ggggtacgtt ctttcccgcc ccgagcccat gtgggcggag 960
 cttaacgcct tggccgccgc ctgggagggg agggtttacc gggcgaggga tcccttgag 1020
 gccttgccgg ggcttgggga ggtgaggggg cttttggcca aggacctggc ggtgctggcc 1080

ctgaggggaag ggattgccct ggcaccgggc gacgacccca tgctcctcgc ctacctcctg 1140
gatccttcca acaccgcccc cgaaggggta gcccggcgct acggggggga gtggaccgag 1200
gaggcggggg aaagggcgct gctttccgaa aggctttacg ccgccctcct gaagcggctt 1260
aagggggagg agaggcttct ttggctttac gaggaggtgg aaaagccctt ttcgcgggtc 1320
ctggcccaca tggaggccac ggggggtacg ttggatgtgg cctacttaaa ggccctttcc 1380
ctggaggtgg aggcggagat aaggcgcttc gaggaggagg tccaccgcct ggccgggcat 1440
cctttcaacc tgaactcccg ggaccagctg gaaagggta tctttgacga gcttgggctt 1500
cccgccatcg gcaagacgca gaagacgggc aagcgctcca ccagcgccgc cgttttgagg 1560
gccttgcggg aggtcatcc catcgtggac cgcctccttc agtaccggga gctttccaag 1620
ctcaagggaa cctacatcga tcccttgctt gccctggtcc accccaagac gaaccgcctc 1680
cacaccggtt tcaaccagac ggccaccgcc acggggaggc ttagcagctc ggatcctaata 1740
ctgcaaaaata tccccgtgcg cacccttttg ggccagcgga tccgccgggc cttcgtggcc 1800
gaggaggggt ggaggtggt ggttttgac tacagccaga ttgagctcag ggtcctggcg 1860
cacctttccg gggacgagaa cctaataccg gtcttccagg agggccagga catccacacc 1920
cagacggcca gctggatgtt cggcgtgccc ccagaggccg tggattccct gatgcgccgg 1980
gcggccaaga ccatcaactt cggcgtcctc tacggcatgt ccgccaccg gctttcggga 2040
gagctggcca tccccctacga ggaggcggtg gccttcatcg agcgggtattt ccagagctac 2100
cccaaggtgc gggcctggat tgagaaaacc ctggcggaag gacgggaacg gggctatgtg 2160
gaaaccctct ttggccgccg gcgctacgtg cccgacttgg cttcccgggt gaagagcatc 2220
cgggaggcag cggagcgcac ggccttcaac atgccgggtc aggggaccgc cgcggatttg 2280
atgaaactgg ccatggtgaa gctctttccc aggttcagg agctgggggc caggatgctt 2340
ttgcaggtgc acaacgaact ggtcctcgag gctcccaagg agcaagcgga ggaagtcgcc 2400
caggaggcca agcggaccat ggaggaggtg tggccccctga aggtgccctt ggaggtggaa 2460
gtgggcatcg gggaggactg gctttccgcc aaggcctag 2499

<210> 132
<211> 2526
<212> DNA
<213> Artificial

<220>
<223> Synthetic

acgggcccgc tccacaccgc cttcaaccag acggccacgg ccacggggag gcttagtagc 1740
 tccgaccca acctgcagaa catccccgtc cgcaccccct tgggccagag gatccgcggg 1800
 gccttcgtgg ccgaggcggg ttgggcgttg gtggccctgg actatagcca gatagagctc 1860
 cgcgtcctcg cccacctctc cggggacgaa aacctgatca gggcttcca ggaggggaag 1920
 gacatccaca cccagaccgc aagctggatg ttcggcgtcc ccccgaggc cgtggacccc 1980
 ctgatgcgcc gggcgccaa gacggtgaac ttcggcgtcc tctacggcat gtccgcccac 2040
 aggtctctcc aggagcttgc catccccctac gaggaggcgg tggcctttat agagcgctac 2100
 ttccaaagct tccccaaggt gcgggcctgg atagaaaaga ccctggagga ggggaggaag 2160
 cggggctacg tggaaaccct cttcggaaga aggcgctacg tgcccacact caacgcccgg 2220
 gtgaagagcg tcaggaggc cgcggagcgc atggccttca acatgcccgt ccagggcacc 2280
 gccgccgacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg 2340
 gcccgcatgc tcctccaggt cgccaacgag ctctccttg agggccccca agcgcggggc 2400
 gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatcccct cgccgtgccc 2460
 ctggaggtgg aggtggggat gggggaggac tggctttccg ccaagggta ccaccaccac 2520
 caccac 2526

<210> 133
 <211> 2514
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 133
 atgaattccc tgcccctctt tgagcccaag ggccgggtgc ttctggtgga cggccaccac 60
 ctggcctacc gtaccttttt tgccctgaag ggcctcacca ccagccgcgg ggagccggtc 120
 caggcgggtg acgggtttgc caagagcctt ttgaaggcgc taagggaaga cggggatgtg 180
 gtgatcgtgg tgtttgacgc caaggcccc tccttcgcc accagacctc cgaggcctac 240
 aaggcggggc gggctccac ccccgaggac tttccccggc agcttgccct tatcaaggag 300
 atggtggacc ttttgggcct ggagcgcctc gaggtgccgg gctttgaagc ggatgacgtc 360
 ctggctaccc tggccaagaa ggcggaaaag gaaggctacg aagtgcgcat cctcaccgcg 420
 gaccgggacc tttaccagct tctttcggag cgaatctcca tccttcaccc ggaggggttac 480

| | | | | | | |
|-------------|-------------|-------------|-------------|------------|-------------|------|
| ctgatcacc | cggagtggct | ttgggagaag | tatgggctta | agccttccca | gtgggtggac | 540 |
| taccgggcct | tggccgggga | cccttccgac | aacatccccg | gcbtgaagg | catcggggag | 600 |
| aagacggcgg | ccaagctgat | ccgggagtgg | ggaagcctgg | aaaaccttct | taagcacctg | 660 |
| gaacaggtga | aacctgcctc | cgtgcgggag | aagatcctta | gccacatgga | ggacctcaag | 720 |
| ctatccctgg | agctatcccg | ggtgcacacg | gacttgctcc | ttcaggtgga | cttcgcccgg | 780 |
| cgccgggagc | cggaccggga | ggggcttaag | gccttttttg | agaggctgga | gttcggaagc | 840 |
| ctcctccacg | agttcggcct | gttgaaaagc | ccggtggcgg | cggaggaagc | tccctggccg | 900 |
| cccccgagg | gagccttcgt | ggggtacgtt | ctttcccgcc | ccgagcccat | gtgggcggag | 960 |
| cttaacgcct | tggccgccc | ctggggcggc | cgcbtgcacc | gggcagcaga | ccccttgggc | 1020 |
| gggctaaagg | acctcaagga | ggtccggggc | ctcctcgcca | aggacctcgc | cgtcttgggc | 1080 |
| tcgagggagg | ggctagacct | cgtgcccggg | gacgacccca | tgctcctcgc | ctacctcctg | 1140 |
| gaccttctga | acaccacccc | cgaggggggtg | gcgcggcgct | acggggggga | gtggacggag | 1200 |
| gacgcgccc | accgggccct | cctctcggag | aggctccatc | ggaacctcct | taagcgccctc | 1260 |
| gagggggagg | agaagctcct | ttggctctac | cacgaggtgg | aaaagcccct | ctcccgggtc | 1320 |
| ctggcccata | tggaggccac | cggggtacgg | cgggacgtgg | cctaccttca | ggccctttcc | 1380 |
| ctggagcttg | cggaggagat | ccgcgcctc | gaggaggagg | tcttccgctt | ggcgggccac | 1440 |
| cccttcaacc | tcaactcccc | ggaccagctg | gaaagggtgc | tctttgacga | gcttaggctt | 1500 |
| ccgccttga | agaagacgaa | gaagacaggc | aagcgtccca | ccagcgccgc | ggtgctggag | 1560 |
| gccttacggg | aggcccaccc | catcgtggag | aagatcctcc | agcaccggga | gctcaccaag | 1620 |
| ctcaagaaca | cctacgtgga | ccccctccca | agcctcgtcc | acctgaggac | gggcgcgcctc | 1680 |
| cacaccgct | tcaaccagac | ggccacggcc | acggggaggc | ttagtagctc | cgaccccaac | 1740 |
| ctgcagaaca | tccccgtccg | cacccccttg | ggccagagga | tccgcccggc | cttcgtggcc | 1800 |
| gaggcgggtt | gggcgttggt | ggccctggac | tatagccaga | tagagctccg | cgtcctcgcc | 1860 |
| cacctctccg | gggacgaaaa | cctgatcagg | gtcttccagg | aggggaagga | catccacacc | 1920 |
| cagaccgcaa | gctggatgtt | cggcgtcccc | ccggaggccg | tggaccccct | gatgcgccgg | 1980 |
| gcggccaaga | cgggtgaactt | cggcgtcctc | tacggcatgt | ccgcccatag | gctctcccag | 2040 |
| gagcttgcca | tcccctacga | ggaggcgggtg | gcctttatag | agcgtactt | ccaaagcttc | 2100 |
| cccaagggtgc | gggcctggat | agaaaagacc | ctggaggagg | ggaggaagcg | gggctacgtg | 2160 |
| gaaacctct | tcggaagaag | gcgctacgtg | cccgcacctca | acgcccgggt | gaagagcgtc | 2220 |

agggaggccg cggagcgcacat ggccttcaac atgcccgtcc agggcaccgc cgccgacctc 2280
 atgaagctcg ccatggtgaa gctcttcccc cgcctccggg agatgggggc ccgcatgctc 2340
 ctccaggtcg ccaacgagct cctcctggag gcccccaag cgcgggccga ggaggtggcg 2400
 gctttggcca aggaggccat ggagaaggcc tatccctcg ccgtgcccct ggaggtggag 2460
 gtggggatgg gggaggactg gctttccgcc aagggtcacc accaccacca ccac 2514

<210> 134
 <211> 2526
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 134
 atgaattcca cccactttt tgacctggag gaacccccca agcgggtgct tctggtggac 60
 ggccaccacc tggcctaccg caccttctat gccctgagcc tcaccacctc ccggggggag 120
 ccggtgcaga tggctctacg cttcgcccgg agcctcctca aggccttgaa ggaggacgga 180
 caggcggtgg tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacgag 240
 gcctacaagg cgggcccggg cccacccccg gaggacttcc cccgccagct cgccttggtc 300
 aagcggtggtg tggaccttct gggcctggtc cgcctcgagg ccccggggta cgaggcggac 360
 gacgtcctgg gcacctggc caagaaggcc gaaaggagg ggatggaggt gcgcatactc 420
 acgggagacc gggacttctt ccagctctc tccgagaagg tctcggtcct cctgccggac 480
 gggaccctgg tcacccccaa ggacgtccag gagaagtacg gggtgcccc ggagcgctgg 540
 gtggacttcc gcgccctcac gggggaccgc tcggacaaca tccccggggt ggcggggata 600
 ggggagaaga ccgcccttcg actcctcgca gagtggggga gcgtggaaaa cctcctgaag 660
 aacctggacc gggtaaagcc ggactcgctc cggcgcaaga tagaggcgca cctcgaggac 720
 ctccacctct ccttagacct ggcccgcac cgcaccgacc tccccctgga ggtggacttt 780
 aaggccctgc gccgcaggac ccccgacctg gagggcctga gggccttttt ggaggagctg 840
 gagttcgga gctcctcca cgagttcggc ctcttgggag gggagaagcc ccgggaggag 900
 gccccctggc ccccgcccga aggggccttc gtgggcttcc tcctttccc caaggagccc 960
 atgtgggcgg agcttctggc cctggcgggc gctcggggc gccgcgtcca ccgggcccc 1020
 gagccttata aagccctcag ggacctgaag gagggcgggg ggcttctcgc caaagacctg 1080

agcgttcttg ccctgagggg aggccttggc ctcccggccg gcgacgaccc catgctcctc 1140
gcctacctcc tggacccttc gaacaccacc cccgaggggg tggcccgggg ctacggcggg 1200
gagtggacgg aggagggcgg ggagcgggcc gccctttccg agaggctctt cgccaacctg 1260
cttaagaggc ttgaggggga ggagaggctc ctttggcttt accgggaggt ggagaggccc 1320
ctttccgctg tcctggccca tatggaggcc acgggggtgc gcctggacgt ggcctatctc 1380
agggccttgt ccctggaggt ggccgaggag atcgcccgcc tcgaggccga ggtcttcgcg 1440
ctggccggcc accccttcaa cctcaactcc cgggaccagc tggaaagggt cctctttgac 1500
gagctagggc ttcccgccat caagaagacg caaaagaccg gcaagcgctc caccagcgcc 1560
gccgtccttg aggccctccg cgaggcccac cccatcgctg agaagatcct gcagtaccgg 1620
gagctcacca agctgaagag cacctacatt gaccccttgc cggacctcat ccaccccgag 1680
acgggcccgc tccacacccg cttcaaccag acggccacgg ccacgggcag gctaagtagc 1740
tccgatccca acctccagaa catccccgtc cgcaccccg c tggggcagag gatccgccgg 1800
gccttcctcg ccgaggaggg gtggtctattg gtggcccttg actatagcca gatagagctc 1860
agggtgctgg cccacctctc cggcgacgag aacctgatcc gggctctcca ggagggcgcg 1920
gacatccaca cggagaccgc cagctggatg ttcggcgctc cccgggaggc cgtggacccc 1980
ctgatgcgcc gggcggccaa gaccatcaac ttcgggggtc tctacggcat gtcggcccac 2040
cgctctctcc aggagctagc catcccttac gaggaggccc aggccttcat tgagcgctac 2100
tttcagagct tccccagggt gcgggccttg attgagaaga ccctggagga gggcaggagg 2160
cgggggtacg tggagaccct cttcggccgc cgccgctacg tgccagacct agaggcccg 2220
gtgaagagcg tgcgggaggc ggccgagcgc atggccttca acatgcccgt ccagggcacc 2280
gccgccgacc tcatgaagct ggctatggtg aagctcttcc ccaggctgga ggaaatgggg 2340
gccaggatgc tccttcaggt cgccaacgag ctggtcctcg agggcccaaa agagagggcg 2400
gaggccgttg cccggctggc caaggaggtc atggaggggg tgtatccctt ggccgtgccc 2460
ctggaggtgg aggtggggat aggggaggac tggtctctcc ccaaggagca ccaccaccac 2520
caccac 2526

<210> 135
<211> 2514
<212> DNA
<213> Artificial
<220>

<223> Synthetic

<400> 135

| | |
|--|------|
| atgaattccc tgccccctctt tgagcccaag ggccgggtgc ttctggtgga cggccaccac | 60 |
| ctggcctacc gtaccttttt tgccctgaag ggccctacca ccagccgcgg ggagccggtc | 120 |
| caggcgggtgt acgggtttgc caagagcctt ttgaaggcgc taagggaaga cggggatgtg | 180 |
| gtgatcgtgg tgtttgacgc caaggcccc tccttcgcc accagaccta cgaggcctac | 240 |
| aaggcggggc gggctccac ccccgaggac tttccccggc agcttgccct tatcaaggag | 300 |
| atggtggacc ttttgggcct ggagcgctc gaggtgccgg gctttgaagc ggatgacgtc | 360 |
| ctggctaccc tggccaagaa ggcgaaaag gaaggctacg aagtgcgcat cctcaccgcg | 420 |
| gaccgggacc tttaccagct tctttcggag cgaatctcca tccttcaccc ggaggggttac | 480 |
| ctgatcacc cggagtggct ttgggagaag tatgggctta agccttccca gtgggtggac | 540 |
| taccgggcct tggccgggga cccttcgac aacatccccg gcgtgaaggg catcggggag | 600 |
| aagacggcgg ccaagctgat ccgggagtgg ggaagcctgg aaaaccttct taagcacctg | 660 |
| gaacaggtga aacctgcctc cgtgcgggag aagatcctta gccacatgga ggacctcaag | 720 |
| ctatccctgg agctatccc ggtgcacacg gacttgctcc ttcagggtgga cttcgcccgg | 780 |
| cgccgggagc cggaccggga ggggcttaag gccttttttg agaggctgga gttcggaagc | 840 |
| ctcctccacg agttcggcct gttggaagc ccggtggcgg cggaggaagc tccctggccg | 900 |
| ccccccgagg gagccttcgt ggggtacgtt ctttccgcc ccgagcccat gtgggcggag | 960 |
| cttaacgcct tggccgcgc ctggggcggc cgcgtccacc gggccccga gccttataaa | 1020 |
| gccctcaggg acctgaagga ggcgcggggg cttctcgcca aagacctgag cgttctggcc | 1080 |
| ctgagggaag gccttggcct ccgcccggc gacgaccca tgctcctcgc ctacctctg | 1140 |
| gacccttcga acaccacccc cgaggggggtg gcccggcgt acggcgggga gtggacggag | 1200 |
| gaggcggggg agcgggcgc cctttccgag aggctcttcg ccaacctgct taagaggctt | 1260 |
| gagggggagg agaggctcct ttggctttac cgggaggtgg agaggccct ttccgctgtc | 1320 |
| ctggcccata tggaggccac gggggtgcgc ctggacgtgg cctatctcag ggccttgctc | 1380 |
| ctggaggtgg ccgaggagat cgcgcgcctc gaggccgagg tcttcgcct ggccggccac | 1440 |
| cccttcaacc tcaactccc ggaccagctg gaaagggtcc tctttgacga gctagggtt | 1500 |
| cccgccatca agaagacgca aaagaccggc aagcgtcca ccagcgccgc cgtcctggag | 1560 |
| gccctccgcg agggccaccc catcgtggag aagatcctgc agtaccggga gctcaccaag | 1620 |

ctgaagagca cctacattga ccccttgccg gacctcatcc accccaggac gggccgcctc 1680
cacacccgct tcaaccagac ggccacggcc acgggcaggc taagtagctc cgatcccaac 1740
ctccagaaca tccccgtccg caccocgctt gggcagagga tccgcccggc cttcatcgcc 1800
gaggaggggt ggctatttgt ggccctggac tatagccaga tagagctcag ggtgctggcc 1860
cacctctccg ggcacgagaa cctgatccgg gtcttccagg aggggcggga catccacacg 1920
gagaccgcca gctggatgtt cggcgtcccc cgggaggccg tggaccccct gatgcgcccg 1980
gcggccaaga ccatcaactt cggggctcctc tacggcatgt cggcccaccg cctctcccag 2040
gagctagcca tcccttacga ggaggcccag gccttcattg agcgctactt tcagagcttc 2100
cccaagggtgc gggcctggat tgagaagacc ctggaggagg gcaggaggcg ggggtacgtg 2160
gagaccctct tcggccgccc ccgctacgtg ccagacctag agggccgggt gaagagcgtg 2220
cgggaggcgg ccgagcgcct ggccctcaac atgcccgtcc agggcaccgc cgcgcacctc 2280
atgaagctgg ctatggtgaa gctcttcccc aggctggagg aaatgggggc caggatgctc 2340
cttcaggctg ccaacgagct ggtcctcgag gccccaaaag agagggcgga ggccgtggcc 2400
cggctggcca aggaggtcat ggaggggggtg tatcccctgg ccgtgcccct ggaggtggag 2460
gtggggatag gggaggactg gctctccgcc aaggagcacc accaccacca ccac 2514

<210> 136
<211> 320
<212> PRT
<213> Artificial

<220>
<223> Synthetic

<400> 136

Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg Ile Asn Ser Gly
1 5 10 15
Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu Val Asp Gly
20 25 30
His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys Gly Leu Thr Thr
35 40 45
Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala Lys Ser Leu
50 55 60
Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile Val Val Phe Asp
65 70 75 80
Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly Gly Tyr Lys Ala

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| | | | | 85 | | | | | | 90 | | | | | 95 |
| Gly | Arg | Ala | Pro | Thr | Leu | Val | Pro | Arg | Gly | Ser | Glu | Asp | Phe | Pro | Arg |
| | | | 100 | | | | | 105 | | | | | 110 | | |
| Gln | Leu | Ala | Leu | Ile | Lys | Glu | Leu | Val | Asp | Leu | Leu | Gly | Leu | Ala | Arg |
| | | 115 | | | | | 120 | | | | | 125 | | | |
| Leu | Glu | Val | Pro | Gly | Tyr | Glu | Ala | Asp | Asp | Val | Leu | Ala | Ser | Leu | Ala |
| | 130 | | | | | 135 | | | | | 140 | | | | |
| Lys | Lys | Ala | Glu | Lys | Glu | Gly | Tyr | Glu | Val | Arg | Ile | Leu | Thr | Ala | Asp |
| 145 | | | | | 150 | | | | | 155 | | | | | 160 |
| Lys | Asp | Leu | Tyr | Gln | Leu | Leu | Ser | Asp | Arg | Ile | His | Val | Leu | His | Pro |
| | | | | 165 | | | | | 170 | | | | | 175 | |
| Glu | Gly | Tyr | Leu | Ile | Thr | Pro | Ala | Trp | Leu | Trp | Glu | Lys | Tyr | Gly | Leu |
| | | | 180 | | | | | 185 | | | | | 190 | | |
| Arg | Pro | Asp | Gln | Trp | Ala | Asp | Tyr | Arg | Ala | Leu | Thr | Gly | Asp | Glu | Ser |
| | | 195 | | | | | 200 | | | | | 205 | | | |
| Asp | Asn | Leu | Pro | Gly | Val | Lys | Gly | Ile | Gly | Glu | Lys | Thr | Ala | Arg | Lys |
| | 210 | | | | | 215 | | | | | 220 | | | | |
| Leu | Leu | Glu | Glu | Trp | Gly | Ser | Leu | Glu | Ala | Leu | Leu | Lys | Asn | Leu | Asp |
| 225 | | | | | 230 | | | | | 235 | | | | | 240 |
| Arg | Leu | Lys | Pro | Ala | Ile | Arg | Glu | Lys | Ile | Leu | Ala | His | Met | Asp | Asp |
| | | | 245 | | | | | | 250 | | | | | 255 | |
| Leu | Lys | Leu | Ser | Trp | Asp | Leu | Ala | Lys | Val | Arg | Thr | Asp | Leu | Pro | Leu |
| | | | 260 | | | | | 265 | | | | | 270 | | |
| Glu | Val | Asp | Phe | Ala | Lys | Arg | Arg | Glu | Pro | Asp | Arg | Glu | Arg | Leu | Arg |
| | | 275 | | | | | 280 | | | | | 285 | | | |
| Ala | Phe | Leu | Glu | Arg | Leu | Glu | Phe | Gly | Ser | Leu | Leu | His | Glu | Phe | Gly |
| | 290 | | | | | 295 | | | | | 300 | | | | |
| Leu | Leu | Glu | Ser | Pro | Lys | Ala | Ala | Leu | Glu | His | His | His | His | His | His |
| 305 | | | | | 310 | | | | | 315 | | | | | 320 |

<210> 137
 <211> 73
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 137
 ccctatcttt aaagttttta aaaagtttga ccccttttg ggggccctat ctttaaagtt 60
 tttaaaaatt tga 73

<210> 138
<211> 15
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 138
cgcgcggaac gcgcg

15

<210> 139
<211> 16
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 139
cccgggtttt cccggg

16

<210> 140
<211> 20
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 140
aggcgcacca atttggtggt

20

<210> 141
<211> 53
<212> RNA
<213> Artificial

<220>
<223> Synthetic

<400> 141
uucgcuuucu ucccuuccuu ucucgccacg uucgccggcu uuccccguca agc

53

<210> 142
<211> 30
<212> DNA
<213> Artificial

<220>
<223> Synthetic

30

| | |
|-------|------------|
| <210> | 143 |
| <211> | 26 |
| <212> | DNA |
| <213> | Artificial |

<220>
<223> Synthetic

```
<400> 143
attagaaagg aagggaagaa agcgaa
```

26

```
<210> 144
<211> 30
<212> DNA
<213> Artificial
```

<220>
<223> Synthetic

```
<400> 144
acggggaaag ccggcgaacg tggcgagaac
```

30

| | |
|-------|------------|
| <210> | 145 |
| <211> | 30 |
| <212> | DNA |
| <213> | Artificial |

<220>
<223> Synthetic

```
<400> 145
cttgacgggg aaagccggcg aacgtggcgc
```

30

| | |
|-------|------------|
| <210> | 146 |
| <211> | 23 |
| <212> | DNA |
| <213> | Artificial |

<220>
<223> Synthetic

```
<400> 146
agaaaggaag ggaagaaagc gaa
```

23

| | |
|-------|------------|
| <210> | 147 |
| <211> | 63 |
| <212> | DNA |
| <213> | Artificial |

```

<220>
<223> Synthetic

<400> 147
gcgcggcggc ggggtgtggtg gttacgcgca gcgtgaccgc tacacttgcc agcgccctag      60

cgc                                                                    63

<210> 148
<211> 30
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 148
gcgctagggc gctggcaagt gtagcgggtca      30

<210> 149
<211> 35
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 149
gatcgctgcg cgtaaccacc acaccgcgc cgcg      35

<210> 150
<211> 30
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 150
ggactctgcc tcaagacggt agtcaacgtg      30

<210> 151
<211> 16
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 151
cacgttgact accgtc      16

```


<210> 152
 <211> 34
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 152
 catgtcaagc agtcctaact ttgaggcaga gtcc 34

<210> 153
 <211> 2506
 <212> DNA
 <213> Thermus aquaticus

<400> 153
 atgaggggga tgctgcccct ctttgagccc aagggccggg tcctcctggt ggacggccac 60
 cacctggcct accgcacctt ccacgccctg aagggcctca ccaccagccg gggggagccg 120
 gtgcaggcgg tctacggctt cgccaagagc ctctcaagg ccctcaagga ggacggggac 180
 gcggtgatcg tgggtctttga cgccaaggcc ccctccttcc gccacgaggc ctacgggggg 240
 tacaaggcgg gccggggccc cagcgggag gactttcccc ggcaactcgc cctcatcaag 300
 gagctggtgg acctcctggg gctggcgcg ctcgaggtcc cgggctacga ggcggacgac 360
 gtctggcca gcctggccaa gaaggcggaa aaggagggt acgaggtccg catcctcacc 420
 gccgacaaag acctttacca gctcctttcc gaccgcatcc acgtcctcca ccccgagggg 480
 tacctcatca ccccggcctg gctttgggaa aagtacggcc tgaggcccg cagtgggcc 540
 gactaccggg ccctgaccgg ggacgagtcc gacaaccttc ccgggggtcaa gggcatcggg 600
 gagaagacgg cgaggaagct tctggaggag tgggggagcc tggaagccct cctcaagaac 660
 ctggaccggc tgaagccgc catccgggag aagatcctgg cccacatgga cgatctgaag 720
 ctctcctggg acctggccaa ggtgcgcacc gacctgcccc tggaggtgga cttcgccaaa 780
 aggcgggagc ccgaccggga gaggcttagg gcctttctgg agaggcttga gtttggcagc 840
 ctctccacg agttcggcct tctggaaagc cccaaggccc tggaggaggc cccctggccc 900
 ccgccggaag gggccttcgt gggctttgtg ctttcccga aggagcccat gtgggccgat 960
 cttctggccc tgcccgccgc cagggggggc cgggtccacc gggccccga gccttataaa 1020
 gccctcaggg acctgaagga ggcgcggggg cttctcgcca aagacctgag cgttctggcc 1080
 ctgagggaa gcttggcct cccgccggc gacgaccca tgctcctcgc ctacctctg 1140
 gaccttcca acaccacccc cgaggggggtg gcccggcgt acggcgggga gtggacggag 1200

gaggcggggg agcggggccgc cctttccgag aggctcttcg ccaacctgtg ggggaggctt 1260
gagggggagg agaggctcct ttggctttac cgggaggtgg agaggcccct ttccgctgtc 1320
ctggcccaca tggaggccac gggggtgcmc ctggacgtgg cctatctcag ggccttgtcc 1380
ctggaggtgg ccgaggagat cggccgcctc gaggccgagg tcttccgcct ggccggccac 1440
cccttcaacc tcaactcccg ggaccagctg gaaagggtcc tctttgacga gctagggtt 1500
cccgccatcg gcaagacgga gaagaccggc aagcgctcca ccagcgccgc cgtcctggag 1560
gccctccgcg agggccaccc catcgtggag aagatcctgc agtaccggga gctcaccaag 1620
ctgaagagca cctacattga ccccttgccg gacctcatcc accccaggac gggccgcctc 1680
cacaccgct tcaaccagac ggccacggcc acgggcaggc taagtagctc cgatcccaac 1740
ctccagaaca tccccgtccg cccccgctt gggcagagga tccgcggggc cttcatcgcc 1800
gaggaggggt ggctattggt ggccctggac tatagccaga tagagctcag ggtgctggcc 1860
cacctctccg gcgacgagaa cctgatccgg gtcttccagg aggggcggga catccacacg 1920
gagaccgcca gctggatggt cggcgctccc cgggaggccg tggacccct gatgcgccg 1980
gcgcccaaga ccatcaactt cggggtcctc tacggcatgt cggcccaccg cctctcccag 2040
gagctagcca tcccttacga ggaggcccag gccttcattg agcgctactt tcagagcttc 2100
cccaagggtgc gggcctggat tgagaagacc ctggaggagg gcaggaggcg ggggtacgtg 2160
gagaccctct tcggccggcg ccgctacgtg ccagacctag agggccgggt gaagagcgtg 2220
cgggaggcgg ccgagcgcat ggccttcaac atgcccgtcc agggcaccgc cgccgacctc 2280
atgaagctgg ctatggtgaa gctcttcccc aggctggagg aaatgggggc caggatgctc 2340
cttcagggtcc acgacgagct ggtcctcgag gcccctaaag agagggcgga ggccgtggcc 2400
cggctggcca aggaggtcat ggaggggggtg tatcccctgg ccgtgcccct ggaggtggag 2460
gtgggggatag gggaggactg gctctccgcc aaggagtgat accacc 2506

<210> 154
<211> 2496
<212> DNA
<213> *Thermus flavus*

<400> 154
atggcgatgc ttcccctctt tgagcccaaa ggccgcgtgc tcctggtgga cggccaccac 60
ctggcctacc gcaccttctt tgccctcaag ggccctacca ccagccgcgg cgaaccggtt 120
caggcggctc acggcttcgc caaaagcctc ctcaaggccc tgaaggagga cggggacgtg 180

| | |
|---|------|
| gtggtggtgg tctttgacgc caaggccccc tcttccgcc acgaggccta cgaggcctac | 240 |
| aaggcggggc gggcccccac cccggaggac tttccccggc agctggccct catcaaggag | 300 |
| ttggtggacc tcctaggcct tgtgcggtg gaggttcccg gctttgaggc ggacgacgtg | 360 |
| ctggccaccc tggccaagcg ggcggaaaag gaggggtacg aggtgcgcat cctcactgcc | 420 |
| gaccgcgacc tctaccagct cctttcggag cgcacgcca tctccaccc tgaggggtac | 480 |
| ctgatcacc cggcgtggct ttacgagaag tacggcctgc gcccgagca gtgggtggac | 540 |
| taccggggcc tggcggggga cccctcgat aacatccccg gggtaaggag catcggggag | 600 |
| aagaccgcc agaggctcat ccgcgagtgg gggagcctgg aaaacctctt ccagcacctg | 660 |
| gaccaggtga agccctcctt gcgggagaag ctccaggcgg gcatggaggc cctggccctt | 720 |
| tcccggaagc tttcccaggt gcacactgac ctgcccctgg aggtggactt cgggaggcgc | 780 |
| cgcacacca acctggaggg tctgcgggct tttttggagc ggttggagtt tggaagcctc | 840 |
| ctccacgagt tcggcctcct ggagggggcg aaggcggcag aggaggcccc ctggccccct | 900 |
| ccggaagggg cttttttggg cttttccttt tcccgcccc agcccatgtg ggccgagctt | 960 |
| ctggccctgg ctggggcgtg ggagggggcg ctccatcggg cacaagaccc ccttaggggc | 1020 |
| ctgagggacc ttaagggggt gcggggaatc ctggccaagg acctggcggg tttggccctg | 1080 |
| cgggagggcc tggacctctt cccagaggac gaccccatgc tcttggccta ccttctggac | 1140 |
| ccctccaaca ccacctctga ggggggtggc cggcgttacg ggggggagtg gacggaggat | 1200 |
| gcgggggaga gggccctcct ggccgagcgc ctcttcaga ccctaaagga gcgccttaag | 1260 |
| ggagaagaac gcctgctttg gctttacgag gaggtggaga agccgctttc ccgggtgttg | 1320 |
| gcccggatgg aggccacggg ggtccggctg gacgtggcct acctccaggc cctctccctg | 1380 |
| gaggtggagg cggaggtgcg ccagctggag gaggaggtct tccgcctggc cgccacccc | 1440 |
| ttcaacctca actcccgcga ccagctggag cgggtgctct ttgacgagct gggcctgcct | 1500 |
| gccatcggca agacggagaa gacggggaaa cgctccacca gcgctgccgt gctggaggcc | 1560 |
| ctgcgagagg cccaccccat cgtggaccgc atcctgcagt accgggagct caccaagctc | 1620 |
| aagaacacct acatagaccc cctgcccgcc ctggtccacc ccaagaccgg ccggctccac | 1680 |
| accgcttca accagacggc caccgccacg ggcaggcttt ccagctccga cccaacctg | 1740 |
| cagaacatcc ccgtgcgcac ccctctgggc cagcgcaccc gccgagcctt cgtggccgag | 1800 |
| gagggctggg tgctggtggt cttggactac agccagattg agcttcgggt cctggcccac | 1860 |

<210> 156
 <211> 2502
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<220>
 <221> n
 <222> (1)..(2502)
 <223> n is any of the bases g, a, t, or c

<400> 156
 atgnnggcga tgcttccctt ctttgagccc aaaggccggg tcttcctggt ggacggccac 60
 cacctggcct accgcacett cttcgccctg aagggcctca ccaccagccg gggcgaaccg 120
 gtgcaggcgg tctacggctt cgccaagagc ctctcaagg ccctgaagga ggacggggac 180
 nnggcggtgn tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcttacgag 240
 gcctacaagg cggggccgggc cccaccccg gaggactttc cccggcagct cgcctcatc 300
 aaggagctgg tggacctctt ggggcttgcg cgcctcgagg tccccggcta cgaggcggac 360
 gacgtinctg ccacctggc caagaaggcg gaaaaggagg ggtacgaggt gcgcatactc 420
 accgcgacc gcgacctcta ccagctcctt tccgaccga tcgccgtcct ccaccccgag 480
 gggtaacctca tcaccccggc gtggcttttg gagaagtacg gcctgaggcc ggagcagtgg 540
 gtggactacc gggccctggc gggggacccc tccgacaacc tccccggggt caagggcatac 600
 ggggagaaga ccgcccngaa gctcctcnag gagtggggga gcctggaaaa cctcctcaag 660
 aacctggacc ggggtgaagc cgccntccgg gagaagatcc aggccacat ggangacctg 720
 angctctcct gggagctntc ccaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780
 aagnggcggg agcccgaccg ggaggggctt agggcctttc tggagaggct ggagtttgcc 840
 agcctcctcc acgagttcgg cctcctggag ggccccaagg ccctggagga ggccccctgg 900
 ccccgcccg aaggggcctt cgtgggcttt gtcttttccc gccccgagcc catgtgggcc 960
 gagcttcttg ccctggccgc cgccaggag ggcgggtcc accgggcacc agaccccttt 1020
 angggcctna gggacctnaa ggaggtgcgg ggnctcctcg ccaaggacct ggccgttttg 1080
 gccctgaggg agggcctnga cctcntgccc ggggacgacc ccatgctcct cgcctacctc 1140
 ctggaccctt ccaacaccac ccccgagggg gtggcccggc gctacggggg ggagtggacg 1200
 gaggangcgg gggagcgggc cctcctntcc gagaggctct tcngaacct nnngcagcgc 1260

cttgaggggg aggagaggct cctttggctt taccaggagg tggagaagcc cctttcccgg 1320
gtcctggccc acatggaggc cacgggggtn cggtggacg tggcctacct ccaggccctn 1380
tccctggagg tggcggagga gatccgccg ctcgaggagg aggtcttccg cctggccggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tgctctttga cgagctnngg 1500
cttcccgccca tcggcaagac ggagaagacn ggcaagcgct ccaccagcgc cgccgtgctg 1560
gaggccctnc gngaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
aagctcaaga acacctacat ngacccccctg ccngncctcg tccaccccag gacggggccg 1680
ctccacaccc gcttcaacca gacggccacg gccacgggca ggcttagtag ctccgacccc 1740
aacctgcaga acatccccgt ccgcacccn ctggggcaga ggatccgccg ggcccttcgtg 1800
gccgaggagg gntgggtgtt ggtggccctg gactatagcc agatagagct ccgggtcctg 1860
gcccacctct ccggggacga gaacctgatc cgggtcttcc aggaggggag ggacatccac 1920
accagaccg ccagctggat gttcggcgtc cccccggagg ccgtggaccc cctgatgcgc 1980
cgggcggcca agaccatcaa cttcggggtc ctctacggca tgtccgccca ccgcctctcc 2040
caggagcttg ccatccocta cgaggaggcg gtggccttca ttgagcgcta cttccagagc 2100
ttccccaagg tgccggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
gtggagaccc tcttcggccg ccggcgctac gtgcccgaac tcaacgcccg ggtgaagagc 2220
gtgcgggagg cggcggagcg catggccttc aacatgcccg tccagggcac cgccccgac 2280
ctcatgaagc tggccatggt gaagctcttc ccccggtnc aggaaatggg ggccaggatg 2340
ctcctncagg tccacgacga gctggctctc gaggccccca aagagcgggc ggagnggtg 2400
gccgctttgg ccaaggaggt catggagggg gtctatcccc tggccgtgcc cctggaggtg 2460
gaggtgggga tgggggagga ctggctctcc gccaaaggat ag 2502

<210> 157

<211> 832

<212> PRT

<213> Thermus aquaticus

<400> 157

Met Arg Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu
1 5 10 15

Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys Gly
20 25 30

Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala

| | | |
|---|-----|-----|
| 340 | 345 | 350 |
| Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu Pro | | |
| 355 | 360 | 365 |
| Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser Asn | | |
| 370 | 375 | 380 |
| Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr Glu | | |
| 385 | 390 | 395 |
| Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn Leu | | |
| 405 | 410 | 415 |
| Trp Gly Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg Glu | | |
| 420 | 425 | 430 |
| Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr Gly | | |
| 435 | 440 | 445 |
| Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val Ala | | |
| 450 | 455 | 460 |
| Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly His | | |
| 465 | 470 | 475 |
| Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe Asp | | |
| 485 | 490 | 495 |
| Glu Leu Gly Leu Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys Arg | | |
| 500 | 505 | 510 |
| Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro Ile | | |
| 515 | 520 | 525 |
| Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser Thr | | |
| 530 | 535 | 540 |
| Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg Leu | | |
| 545 | 550 | 555 |
| His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser Ser | | |
| 565 | 570 | 575 |
| Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly Gln | | |
| 580 | 585 | 590 |
| Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val Ala | | |
| 595 | 600 | 605 |
| Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser Gly | | |
| 610 | 615 | 620 |
| Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His Thr | | |
| 625 | 630 | 635 |
| Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp Pro | | |
| 645 | 650 | 655 |

| | 85 | | | | | | | 90 | | | | | | 95 | | | | | |
|---|----|--|--|--|--|--|--|----|--|--|--|--|--|----|--|--|--|--|--|
| Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Val Arg Leu Glu Val 100 105 110 | | | | | | | | | | | | | | | | | | | |
| Pro Gly Phe Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys Arg Ala 115 120 125 | | | | | | | | | | | | | | | | | | | |
| Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Arg Asp Leu 130 135 140 | | | | | | | | | | | | | | | | | | | |
| Tyr Gln Leu Leu Ser Glu Arg Ile Ala Ile Leu His Pro Glu Gly Tyr 145 150 155 160 | | | | | | | | | | | | | | | | | | | |
| Leu Ile Thr Pro Ala Trp Leu Tyr Glu Lys Tyr Gly Leu Arg Pro Glu 165 170 175 | | | | | | | | | | | | | | | | | | | |
| Gln Trp Val Asp Tyr Arg Ala Leu Ala Gly Asp Pro Ser Asp Asn Ile 180 185 190 | | | | | | | | | | | | | | | | | | | |
| Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Gln Arg Leu Ile Arg 195 200 205 | | | | | | | | | | | | | | | | | | | |
| Glu Trp Gly Ser Leu Glu Asn Leu Phe Gln His Leu Asp Gln Val Lys 210 215 220 | | | | | | | | | | | | | | | | | | | |
| Pro Ser Leu Arg Glu Lys Leu Gln Ala Gly Met Glu Ala Leu Ala Leu 225 230 235 240 | | | | | | | | | | | | | | | | | | | |
| Ser Arg Lys Leu Ser Gln Val His Thr Asp Leu Pro Leu Glu Val Asp 245 250 255 | | | | | | | | | | | | | | | | | | | |
| Phe Gly Arg Arg Arg Thr Pro Asn Leu Glu Gly Leu Arg Ala Phe Leu 260 265 270 | | | | | | | | | | | | | | | | | | | |
| Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu Leu Glu 275 280 285 | | | | | | | | | | | | | | | | | | | |
| Gly Pro Lys Ala Ala Glu Glu Ala Pro Trp Pro Pro Pro Glu Gly Ala 290 295 300 | | | | | | | | | | | | | | | | | | | |
| Phe Leu Gly Phe Ser Phe Ser Arg Pro Glu Pro Met Trp Ala Glu Leu 305 310 315 320 | | | | | | | | | | | | | | | | | | | |
| Leu Ala Leu Ala Gly Ala Trp Glu Gly Arg Leu His Arg Ala Gln Asp 325 330 335 | | | | | | | | | | | | | | | | | | | |
| Pro Leu Arg Gly Leu Arg Asp Leu Lys Gly Val Arg Gly Ile Leu Ala 340 345 350 | | | | | | | | | | | | | | | | | | | |
| Lys Asp Leu Ala Val Leu Ala Leu Arg Glu Gly Leu Asp Leu Phe Pro 355 360 365 | | | | | | | | | | | | | | | | | | | |
| Glu Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser Asn Thr 370 375 380 | | | | | | | | | | | | | | | | | | | |
| Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr Glu Asp | | | | | | | | | | | | | | | | | | | |

| | | | | | | |
|---|-----------------------------|---------------------|---------------------|-----|--|-----|
| 385 | | 390 | | 395 | | 400 |
| Ala Gly Glu Arg | Ala Leu Leu Ala Glu Arg | Leu Phe Gln Thr | Leu Lys | | | |
| | 405 | 410 | 415 | | | |
| Glu Arg Leu Lys | Gly Glu Glu Arg | Leu Leu Trp | Leu Tyr Glu Glu Val | | | |
| | 420 | 425 | 430 | | | |
| Glu Lys Pro Leu Ser Arg Val | Leu Ala Arg Met | Glu Ala Thr Gly Val | | | | |
| | 435 | 440 | 445 | | | |
| Arg Leu Asp Val Ala Tyr | Leu Gln Ala Leu Ser | Leu Glu Val Glu Ala | | | | |
| | 450 | 455 | 460 | | | |
| Glu Val Arg Gln Leu Glu Glu Glu Val | Phe Arg Leu Ala Gly His Pro | | | | | |
| | 465 | 470 | 475 | | | 480 |
| Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val | Leu Phe Asp Glu | | | | | |
| | 485 | 490 | 495 | | | |
| Leu Gly Leu Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys Arg Ser | | | | | | |
| | 500 | 505 | 510 | | | |
| Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro Ile Val | | | | | | |
| | 515 | 520 | 525 | | | |
| Asp Arg Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Asn Thr Tyr | | | | | | |
| | 530 | 535 | 540 | | | |
| Ile Asp Pro Leu Pro Ala Leu Val His Pro Lys Thr Gly Arg Leu His | | | | | | |
| | 545 | 550 | 555 | | | 560 |
| Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser Ser Ser | | | | | | |
| | 565 | 570 | 575 | | | |
| Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly Gln Arg | | | | | | |
| | 580 | 585 | 590 | | | |
| Ile Arg Arg Ala Phe Val Ala Glu Glu Gly Trp Val Leu Val Val Leu | | | | | | |
| | 595 | 600 | 605 | | | |
| Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser Gly Asp | | | | | | |
| | 610 | 615 | 620 | | | |
| Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His Thr Gln | | | | | | |
| | 625 | 630 | 635 | | | 640 |
| Thr Ala Ser Trp Met Phe Gly Val Ser Pro Glu Gly Val Asp Pro Leu | | | | | | |
| | 645 | 650 | 655 | | | |
| Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr Gly Met | | | | | | |
| | 660 | 665 | 670 | | | |
| Ser Ala His Arg Leu Ser Gly Glu Leu Ser Ile Pro Tyr Glu Glu Ala | | | | | | |
| | 675 | 680 | 685 | | | |
| Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Tyr Pro Lys Val Arg Ala | | | | | | |

700

Glu Val Glu Val Gly Leu Gly Glu Asp Trp Leu Ser Ala Lys Glu
820 825 830

<213> Artificial

<223> Synthetic

<223> X is any amino acid

<400> 159

Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu Ala
65 70 75 80

Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu
 85 90 95
 Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Xaa Arg Leu Glu
 100 105 110
 Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys Lys
 115 120 125
 Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Arg Asp
 130 135 140
 Leu Tyr Gln Leu Leu Ser Asp Arg Ile Ala Val Leu His Pro Glu Gly
 145 150 155 160
 Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg Pro
 165 170 175
 Glu Gln Trp Val Asp Tyr Arg Ala Leu Xaa Gly Asp Pro Ser Asp Asn
 180 185 190
 Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Xaa Lys Leu Leu
 195 200 205
 Xaa Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu Asp Arg Val
 210 215 220
 Lys Pro Xaa Xaa Arg Glu Lys Ile Xaa Ala His Met Glu Asp Leu Xaa
 225 230 235 240
 Leu Ser Xaa Xaa Leu Ser Xaa Val Arg Thr Asp Leu Pro Leu Glu Val
 245 250 255
 Asp Phe Ala Xaa Arg Arg Glu Pro Asp Arg Glu Gly Leu Arg Ala Phe
 260 265 270
 Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu Leu
 275 280 285
 Glu Xaa Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu Gly
 290 295 300
 Ala Phe Val Gly Phe Val Leu Ser Arg Pro Glu Pro Met Trp Ala Glu
 305 310 315 320
 Leu Leu Ala Leu Ala Ala Ala Arg Xaa Gly Arg Val His Arg Ala Xaa
 325 330 335
 Asp Pro Leu Xaa Gly Leu Arg Asp Leu Lys Glu Val Arg Gly Leu Leu
 340 345 350
 Ala Lys Asp Leu Ala Val Leu Ala Leu Arg Glu Gly Leu Asp Leu Xaa
 355 360 365
 Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser Asn
 370 375 380

Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val Arg
690 695 700

Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr Val
705 710 715 720

Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn Ala Arg
725 730 735

Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro
740 745 750

Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu
755 760 765

Phe Pro Arg Leu Xaa Glu Met Gly Ala Arg Met Leu Leu Gln Val His
770 775 780

Asp Glu Leu Val Leu Glu Ala Pro Lys Xaa Arg Ala Glu Xaa Val Ala
785 790 795 800

Ala Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val Pro
805 810 815

Leu Glu Val Glu Val Gly Xaa Gly Glu Asp Trp Leu Ser Ala Lys Glu
820 825 830

Xaa

<210> 160
<211> 640
<212> RNA
<213> Artificial

<220>
<223> Synthetic

<400> 160
gggagcccag cuaugaacuc cuucuccaca agcgccuucg guccaguugc cuucuccug 60
gggcugcucc ugguguugcc ugcugccuuc ccugccccag uacccccagg agaagauucc 120
aaagauguag ccgccccaca cagacagcca cucaccucuu cagaacgaau ugacaaacaa 180
auucgguaa uccucgacgg caucucagcc cugagaaagg agacauguaa caagaguaac 240
augugugaaa gcagcaaaga ggcacuggca gaaaacaacc ugaaccuucc aaagauggcu 300
gaaaaaagau gaugcuucca aucuggauuc aaugaggaga cuugccuggu gaaaaucauc 360
acuggucuuu uggaguuuuga gguauaccua gaguaccucc agaacagauu ugagaguagu 420
gaggaacaag ccagagcugu ccagaugagu aaaaagucc ugauccaguu ccugcagaaa 480
aaggcaaaga aucuagaugc aaauaccacc ccugacccaa ccacaaaugc cagccugcug 540

Asp Pro Ser Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly
 1 5 10 15
 Glu Trp Thr Glu Asp Ala Ala His Arg Ala Leu Leu Ser Glu Arg Leu
 20 25 30
 His Arg Asn Leu Leu Lys Arg Leu Glu Gly Glu Glu Lys Leu Leu Trp
 35 40 45
 Leu Tyr His Glu Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met
 50 55 60
 Glu Ala Thr Gly Val Arg Arg Asp Val Ala Tyr Leu Gln Ala Leu Ser
 65 70 75 80
 Leu Glu Leu Ala Glu Glu Ile Arg Arg Leu Glu Glu Glu Val Phe Arg
 85 90 95
 Leu Ala Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg
 100 105 110
 Val Leu Phe Asp Glu Leu Arg Leu Pro Ala Leu Gly Lys Thr Gln Lys
 115 120 125
 Thr Gly Lys Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu
 130 135 140
 Ala His Pro Ile Val Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys
 145 150 155 160
 Leu Lys Asn Thr Tyr Val Asp Pro Leu Pro Ser Leu Val His Pro Arg
 165 170 175
 Thr Gly Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly
 180 185 190
 Arg Leu Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr
 195 200 205
 Pro Leu Gly Gln Arg Ile
 210

<210> 166

<211> 37

<212> RNA

<213> Artificial

<220>

<223> Synthetic

<220>

<221> n

<222> (1)..(1)

<223> TH 5' end is modified with a biotin-streptavitin complex.

<400> 166
nucacggcag uuggugcgcc ucggaacgag gcgcacg 37

<210> 167
<211> 17
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<220>
<221> n
<222> (1)..(1)
<223> TH 5' end is labeled with tetrachlororfluorescein.

<400> 167
nttttcaact gccgtga 17

<210> 168
<211> 36
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 168
tcacggcagt tggtgcgctt cggaacgagg cgcacg 36

<210> 169
<211> 30
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 169
cgccgagatc acctttacat tttctatcgt 30

<210> 170
<211> 24
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 170
ccttccttat cctggatcctt ggca 24

<210> 171
<211> 24
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 171
acgatagaaa atgtaaaggt gatc

24

<210> 172
<211> 15
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<220>
<221> n
<222> (4)..(4)
<223> The residue at this position is a z28 linker group.

<400> 172
ctcntttctca gtgcg

15

<210> 173
<211> 29
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 173
cgcagtgaga atgaggtgat ctcggcggt

29

<210> 174
<211> 31
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 174
ccgccgagat cacggatggt gtaatcagag a

31

<210> 175
<211> 21

<212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 175 21
 gtgcagggtt gactccttct c

 <210> 176
 <211> 21
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 176 21
 gtgcagggtt gactctttct c

 <210> 177
 <211> 21
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 177 21
 gtgcagggtc gactctttct c

 <210> 178
 <211> 25
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 178 25
 tctctgatta caacatccgt gatct

 <210> 179
 <211> 29
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 179 29
 cgccgagatc acgtagttga ggtcaatga

<400> 184
cagcctcaag attaccgtga tct

23

<210> 185
<211> 23
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 185
ccgtcacgcc tcctccacgg ctc

23

<210> 186
<211> 23
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 186
aggcgaaagc cctcaatttc cca

23

<210> 187
<211> 15
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 187
aaccactgcc gcaca

15

<210> 188
<211> 17
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 188
gagccgtgga ggaggcg

17

<210> 189
<211> 14
<212> DNA
<213> Artificial

<220>
 <223> Synthetic

<220>
 <221> n
 <222> (4)..(4)
 <223> The residue at this position is a z28 linker group.

<400> 189
 cacntgcttc gtgg 14

<210> 190
 <211> 27
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 190
 ccaggaagca agtggaggcg tgacggt 27

<210> 191
 <211> 26
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 191
 ccgtcacgcc tccttcggag tttggg 26

<210> 192
 <211> 25
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 192
 gggttgtgga gtgagtgttc aagta 25

<210> 193
 <211> 20
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

| | |
|--|----|
| <400> 193 gggaaactcc gaaggaggcg | 20 |
| <210> 194 <211> 23 <212> DNA <213> Artificial | |
| <220> <223> Synthetic | |
| <400> 194 ccgtcacgcc tctctgactg cca | 23 |
| <210> 195 <211> 27 <212> DNA <213> Artificial | |
| <220> <223> Synthetic | |
| <400> 195 ttgtcactcg gggttcgaga agatgaa | 27 |
| <210> 196 <211> 11 <212> DNA <213> Artificial | |
| <220> <223> Synthetic | |
| <400> 196 gggccagagg g | 11 |
| <210> 197 <211> 17 <212> DNA <213> Artificial | |
| <220> <223> Synthetic | |
| <400> 197 aggcagtcag agaggcg | 17 |
| <210> 198 <211> 26 <212> DNA <213> Artificial | |

<220>
 <223> Synthetic

 <400> 198
 ccgtcacgcc tcctcctcat tgaatt 26

<210> 199
 <211> 35
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

 <400> 199
 ccaaaagtcc agtgatgatt ttcaccaggc aagta 35

<210> 200
 <211> 20
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

 <400> 200
 cagattggaa gcatccatct 20

<210> 201
 <211> 19
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

 <400> 201
 gattcaatga ggaggaggc 19

<210> 202
 <211> 24
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

 <400> 202
 ccgtcacgcc tccatctggt tagg 24

<210> 203
 <211> 22

<212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 203
 caggtcctgg aaggagcact ta 22

 <210> 204
 <211> 27
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 204
 gccatcagct tctttgttct tgatcatc 27

 <210> 205
 <211> 20
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 205
 gccctaaaca gatggaggcg 20

 <210> 206
 <211> 24
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 206
 ccgtcacgcc tctccagtt gtag 24

 <210> 207
 <211> 30
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 207
 aaaatcatct gtaaattccag cagtaaattga 30


```

<220>
<221> modified_base
<222> (1)..(1)
<223> The residue at tH position contains a TET-label.

<400> 212
attagaaagg aaggaagaa agcgaa 26

<210> 213
<211> 30
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 213
gcttgacggg gaaagccggc gaacgtggcg 30

<210> 214
<211> 30
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 214
cttgacgggg aaagccggcg aacgtggcga 30

<210> 215
<211> 30
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 215
tgacggggaa agccggcgaa cgtggcgaga 30

<210> 216
<211> 30
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 216
acgggggaaag ccggcgaaacg tggcgagaaa 30

```


<212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 221

```

Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
1          5          10          15

Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu
20          25          30

Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
35          40          45

Phe Ala Lys Ser Leu Leu Lys Ala Leu Arg Glu Asp Gly Asp Ala Val
50          55          60

Ile Val Val Phe Asp Ala Glu Ala Pro Ser Phe Arg His Glu Ala Tyr
65          70          75          80

Gly Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg
85          90          95

Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr Arg
100         105         110

Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu Ala
115         120         125

Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp
130         135         140

Lys Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro
145         150         155         160

Glu Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu
165         170         175

Arg Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser
180         185         190

Asp Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu Lys
195         200         205

Leu Leu Lys Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp
210         215         220

Arg Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp
225         230         235         240

Leu Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu
245         250         255

Glu Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Gly Leu Lys

```


[illegible][illegible][illegible][illegible]

| | | |
|-------|-------------------------------------|----|
| <210> | 224 | |
| <211> | 18 | |
| <212> | DNA | |
| <213> | Artificial | |
| <220> | | |
| <223> | Synthetic | |
| <400> | 224 | 18 |
| | acggaacgag cgtctttg | |
| <210> | 225 | |
| <211> | 32 | |
| <212> | RNA | |
| <213> | Artificial | |
| <220> | | |
| <223> | Synthetic | |
| <400> | 225 | 32 |
| | gcgagcgaga cagcgaaaga cgcucguucc gu | |
| <210> | 226 | |
| <211> | 32 | |
| <212> | DNA | |
| <213> | Artificial | |
| <220> | | |
| <223> | Synthetic | |
| <400> | 226 | 32 |
| | gcgagcgaga cagcgaaaga cgctcgttcc gt | |
| <210> | 227 | |
| <211> | 29 | |
| <212> | DNA | |
| <213> | Artificial | |
| <220> | | |
| <223> | Synthetic | |
| <400> | 227 | 29 |
| | acggaacgag cgtctttcat ctgtcaatc | |
| <210> | 228 | |
| <211> | 26 | |
| <212> | RNA | |
| <213> | Artificial | |
| <220> | | |
| <223> | Synthetic | |
| <400> | 228 | |

ucacggcagu uggugcggaa cgcacg

26

<210> 229
<211> 26
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 229
tcacggcagt tggcgcgaa cgcacg

26

<210> 230
<211> 30
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<220>
<221> n
<222> (30)..(30)
<223> TH 3' end is modified with an amine moiety.

<400> 230
cggaggaagc agttggtgcg cctcgttaan

30

<210> 231
<211> 23
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<220>
<221> n
<222> (1)..(1)
<223> TH 5' end is labeled with fluorescein.

<400> 231
ntccttctca actgcttcct ccg

23

<210> 232
<211> 28
<212> DNA
<213> Artificial

<220>

<223> Synthetic

<220>

<221> n

<222> (28)..(28)

<223> TH 3' end is modified with a biotin moiety.

<400> 232

aacgaggcgc acctcaaadc tccctttt

28

<210> 233

<211> 26

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<220>

<221> n

<222> (1)..(1)

<223> TH 5' end is labeled with fluorescein.

<400> 233

nagcgagaca gcgaaagacg ctcggt

26

<210> 234

<211> 17

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<220>

<221> n

<222> (1)..(1)

<223> TH 5' end is labeled with fluorescein.

<400> 234

nttttcgctg tctcgct

17

<210> 235

<211> 13

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 235

| | |
|---|------|
| aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggccacat ggacgatctg | 720 |
| aagctctcct gggacctggc caaggtgcmc accgacctgc ccctggaggt ggacttcgcc | 780 |
| aaaaggcggg agccccgacc ggagaggctt agggcctttc tggagaggct tgagtttggc | 840 |
| agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg | 900 |
| ccccgcggg aaggggcctt cgtgggcctt gtgctttccc gcaaggagcc catgtgggccc | 960 |
| gatctttctg ccctggccgc cgccaggggg ggccgggtcc accgggcccc cgagccttat | 1020 |
| aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg | 1080 |
| gccctgaggg aaggccttgg cctccgcgcc ggcgacgacc ccatgctcct cgctacctc | 1140 |
| ctggaccctt ccaacaccac ccccgagggg gtggccccgc gctacggcgg ggagtggacg | 1200 |
| gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gtgggggagg | 1260 |
| cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct | 1320 |
| gtcctggccc acatggaggc cacgggggtg cgcttgagc tggcctatct cagggccttg | 1380 |
| tccttgagg tggccgggga gatcggccgc ctcgaggccg aggtcttccg cctggccggc | 1440 |
| cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg | 1500 |
| cttcccgcca tcggcaagac ggagaagacc ggcaagcgt ccaccagcgc cgccgtcctg | 1560 |
| gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc | 1620 |
| aagctgaaga gcacctacat tgacccttg ccggacctca tccaccccag gacgggccgc | 1680 |
| ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc | 1740 |
| aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccc ggccttcac | 1800 |
| gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg | 1860 |
| gcccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac | 1920 |
| acggagaccg ccagctggat gttcggcgtc ccccgggagg ccgtggaccc cctgatgcgc | 1980 |
| cgggcgcca agaccatcaa cttcggggtc ctctacggca tgtcggccca ccgcctctcc | 2040 |
| caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc | 2100 |
| ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac | 2160 |
| gtggagaccc tcttcggccg ccgccgtac gtgccagacc tagaggcccg ggtgaagagc | 2220 |
| gtgcgggagg cggccgagcg catggccttc aacatgcccg tccggggcac cgccgccgac | 2280 |
| ctcatgaagc tggctatggt gaagctcttc ccagggctgg aggaaatggg ggccaggatg | 2340 |

ctccttcagg tccacgacga gctggctctc gagggcccaa aagagagggc ggaggccgtg 2400
 gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
 gaggtgggga taggggagga ctggctctcc gccaaaggag ga 2502

<210> 239
 <211> 833
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 239

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Asn | Ser | Gly | Met | Leu | Pro | Leu | Phe | Glu | Pro | Lys | Gly | Arg | Val | Leu |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Leu | Val | Asp | Gly | His | His | Leu | Ala | Tyr | Arg | Thr | Phe | His | Ala | Leu | Lys |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Gly | Leu | Thr | Thr | Ser | Arg | Gly | Glu | Pro | Val | Gln | Ala | Val | Tyr | Gly | Phe |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Ala | Lys | Ser | Leu | Leu | Lys | Ala | Leu | Lys | Glu | Asp | Gly | Asp | Ala | Val | Ile |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| Val | Val | Phe | Asp | Ala | Lys | Ala | Pro | Ser | Phe | Arg | His | Glu | Ala | Tyr | Gly |
| 65 | | | | | 70 | | | | | 75 | | | | | 80 |
| Gly | Tyr | Lys | Ala | Gly | Arg | Ala | Pro | Thr | Pro | Glu | Asp | Phe | Pro | Arg | Gln |
| | | | | 85 | | | | | 90 | | | | | 95 | |
| Leu | Ala | Leu | Ile | Lys | Glu | Leu | Val | Asp | Leu | Leu | Gly | Leu | Ala | Arg | Leu |
| | | | 100 | | | | | 105 | | | | | | 110 | |
| Glu | Val | Pro | Gly | Tyr | Glu | Ala | Asp | Asp | Val | Leu | Ala | Ser | Leu | Ala | Lys |
| | | 115 | | | | | 120 | | | | | 125 | | | |
| Lys | Ala | Glu | Lys | Glu | Gly | Tyr | Glu | Val | Arg | Ile | Leu | Thr | Ala | Asp | Lys |
| | 130 | | | | | 135 | | | | | 140 | | | | |
| Asp | Leu | Tyr | Gln | Leu | Leu | Ser | Asp | Arg | Ile | His | Val | Leu | His | Pro | Glu |
| 145 | | | | 150 | | | | | | 155 | | | | | 160 |
| Gly | Tyr | Leu | Ile | Thr | Pro | Ala | Trp | Leu | Trp | Glu | Lys | Tyr | Gly | Leu | Arg |
| | | | 165 | | | | | 170 | | | | | | 175 | |
| Pro | Asp | Gln | Trp | Ala | Asp | Tyr | Arg | Ala | Leu | Thr | Gly | Asp | Glu | Ser | Asp |
| | | 180 | | | | | | 185 | | | | | 190 | | |
| Asn | Leu | Pro | Gly | Val | Lys | Gly | Ile | Gly | Glu | Lys | Thr | Ala | Arg | Lys | Leu |
| | 195 | | | | | 200 | | | | | | 205 | | | |
| Leu | Glu | Glu | Trp | Gly | Ser | Leu | Glu | Ala | Leu | Leu | Lys | Asn | Leu | Asp | Arg |

| | | |
|---|-----|---------|
| 210 | 215 | 220 |
| Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu | | |
| 225 | 230 | 235 240 |
| Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu | | |
| | 245 | 250 255 |
| Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala | | |
| | 260 | 265 270 |
| Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu | | |
| | 275 | 280 285 |
| Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu | | |
| | 290 | 295 300 |
| Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala | | |
| 305 | 310 | 315 320 |
| Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala | | |
| | 325 | 330 335 |
| Pro Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu | | |
| | 340 | 345 350 |
| Leu Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu | | |
| | 355 | 360 365 |
| Pro Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser | | |
| | 370 | 375 380 |
| Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr | | |
| 385 | 390 | 395 400 |
| Glu Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn | | |
| | 405 | 410 415 |
| Leu Trp Gly Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg | | |
| | 420 | 425 430 |
| Glu Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr | | |
| | 435 | 440 445 |
| Gly Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val | | |
| | 450 | 455 460 |
| Ala Gly Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly | | |
| 465 | 470 | 475 480 |
| His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe | | |
| | 485 | 490 495 |
| Asp Glu Leu Gly Leu Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys | | |
| | 500 | 505 510 |
| Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro | | |

| | | | | |
|---|--|-----|--|-----|
| 515 | | 520 | | 525 |
| Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser | | | | |
| 530 | | 535 | | 540 |
| Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg | | | | |
| 545 | | 550 | | 555 |
| | | | | 560 |
| Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser | | | | |
| | | 565 | | 570 |
| | | | | 575 |
| Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly | | | | |
| | | 580 | | 585 |
| | | | | 590 |
| Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val | | | | |
| | | 595 | | 600 |
| | | | | 605 |
| Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser | | | | |
| | | 610 | | 615 |
| | | | | 620 |
| Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His | | | | |
| | | 625 | | 630 |
| | | | | 635 |
| | | | | 640 |
| Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp | | | | |
| | | 645 | | 650 |
| | | | | 655 |
| Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr | | | | |
| | | 660 | | 665 |
| | | | | 670 |
| Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu | | | | |
| | | 675 | | 680 |
| | | | | 685 |
| Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val | | | | |
| | | 690 | | 695 |
| | | | | 700 |
| Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr | | | | |
| | | 705 | | 710 |
| | | | | 715 |
| | | | | 720 |
| Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala | | | | |
| | | 725 | | 730 |
| | | | | 735 |
| Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met | | | | |
| | | 740 | | 745 |
| | | | | 750 |
| Pro Val Arg Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys | | | | |
| | | 755 | | 760 |
| | | | | 765 |
| Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val | | | | |
| | | 770 | | 775 |
| | | | | 780 |
| His Asp Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val | | | | |
| | | 785 | | 790 |
| | | | | 795 |
| | | | | 800 |
| Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val | | | | |
| | | 805 | | 810 |
| | | | | 815 |
| Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys | | | | |

820

825

830

Glu

<210> 240

<211> 28

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 240

cacgaattcc gaggcgatgc ttccgctc

28

<210> 241

<211> 30

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 241

tgcagctcga ctaacccttg gcggaaagcc

30

<210> 242

<211> 23

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 242

gcatcgctc ggaattcatg gtc

23

<210> 243

<211> 836

<212> PRT

<213> Thermus thermophilus

<400> 243

Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
1 5 10 15

Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu
20 25 30

Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
35 40 45

| | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| Phe | Ala | Lys | Ser | Leu | Leu | Lys | Ala | Leu | Lys | Glu | Asp | Gly | Tyr | Lys | Ala | |
| 50 | | | | | | 55 | | | | 60 | | | | | | |
| Val | Phe | Val | Val | Phe | Asp | Ala | Lys | Ala | Pro | Ser | Phe | Arg | His | Glu | Ala | |
| 65 | | | | 70 | | | | | 75 | | | | | 80 | | |
| Tyr | Glu | Ala | Tyr | Lys | Ala | Gly | Arg | Ala | Pro | Thr | Pro | Glu | Asp | Phe | Pro | |
| | | | 85 | | | | | 90 | | | | | | 95 | | |
| Arg | Gln | Leu | Ala | Leu | Ile | Lys | Glu | Leu | Val | Asp | Leu | Leu | Gly | Phe | Thr | |
| | | 100 | | | | | | 105 | | | | | 110 | | | |
| Arg | Leu | Glu | Val | Pro | Gly | Tyr | Glu | Ala | Asp | Asp | Val | Leu | Ala | Thr | Leu | |
| | 115 | | | | | | 120 | | | | | 125 | | | | |
| Ala | Lys | Lys | Ala | Glu | Lys | Glu | Gly | Tyr | Glu | Val | Arg | Ile | Leu | Thr | Ala | |
| 130 | | | | | | 135 | | | | | 140 | | | | | |
| Asp | Arg | Asp | Leu | Tyr | Gln | Leu | Val | Ser | Asp | Arg | Val | Ala | Val | Leu | His | |
| 145 | | | | | 150 | | | | | 155 | | | | | 160 | |
| Pro | Glu | Gly | His | Leu | Ile | Thr | Pro | Glu | Trp | Leu | Trp | Glu | Lys | Tyr | Gly | |
| | | | | 165 | | | | | 170 | | | | | 175 | | |
| Leu | Arg | Pro | Glu | Gln | Trp | Val | Asp | Phe | Arg | Ala | Leu | Val | Gly | Asp | Pro | |
| | | 180 | | | | | | 185 | | | | | 190 | | | |
| Ser | Asp | Asn | Leu | Pro | Gly | Val | Lys | Gly | Ile | Gly | Glu | Lys | Thr | Ala | Leu | |
| | 195 | | | | | | 200 | | | | | 205 | | | | |
| Lys | Leu | Leu | Lys | Glu | Trp | Gly | Ser | Leu | Glu | Asn | Leu | Leu | Lys | Asn | Leu | |
| | 210 | | | | | 215 | | | | | 220 | | | | | |
| Asp | Arg | Val | Lys | Pro | Glu | Asn | Val | Arg | Glu | Lys | Ile | Lys | Ala | His | Leu | |
| 225 | | | | | 230 | | | | | 235 | | | | | 240 | |
| Glu | Asp | Leu | Arg | Leu | Ser | Leu | Glu | Leu | Ser | Arg | Val | Arg | Thr | Asp | Leu | |
| | | | | 245 | | | | | 250 | | | | | 255 | | |
| Pro | Leu | Glu | Val | Asp | Leu | Ala | Gln | Gly | Arg | Glu | Pro | Asp | Arg | Glu | Gly | |
| | | 260 | | | | | | 265 | | | | | 270 | | | |
| Leu | Arg | Ala | Phe | Leu | Glu | Arg | Leu | Glu | Phe | Gly | Ser | Leu | Leu | His | Glu | |
| | | 275 | | | | | 280 | | | | | 285 | | | | |
| Phe | Gly | Leu | Leu | Glu | Ala | Pro | Ala | Pro | Leu | Glu | Glu | Ala | Pro | Trp | Pro | |
| | 290 | | | | | 295 | | | | | 300 | | | | | |
| Pro | Pro | Glu | Gly | Ala | Phe | Val | Gly | Phe | Val | Leu | Ser | Arg | Pro | Glu | Pro | |
| 305 | | | | | 310 | | | | | 315 | | | | 320 | | |
| Met | Trp | Ala | Glu | Leu | Lys | Ala | Leu | Ala | Ala | Cys | Arg | Asp | Gly | Arg | Val | |
| | | | | 325 | | | | | 330 | | | | | 335 | | |
| His | Arg | Ala | Ala | Asp | Pro | Leu | Ala | Gly | Leu | Lys | Asp | Leu | Lys | Glu | Val | |
| | | | 340 | | | | | 345 | | | | | 350 | | | |

Ala Val Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly
660 665 670

Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile
675 680 685

Pro Tyr Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe
690 695 700

Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys
705 710 715 720

Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp
725 730 735

Leu Asn Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala
740 745 750

Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala
755 760 765

Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu
770 775 780

Leu Gln Val His Asp Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala
785 790 795 800

Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro
805 810 815

Leu Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu
820 825 830

Ser Ala Lys Gly
835

<210> 244
<211> 2511
<212> DNA
<213> Thermus thermophilus

<400> 244
atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
ggccaccacc tgacctaccg caccttcttc gccctgaagg gctcaccac gagccggggc 120
gaaccggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac 180
gggtacaagg ccgtcttcgt ggtctttgac gccaaaggccc cctccttccg ccacgaggcc 240
tacgaggcct acaaggcggg gagggccccg acccccagag acttcccccg gcagctcgcc 300
ctcatcaagg agctggtgga cctcctgggg ttacccgcc tcgaggtccc cggctacgag 360
gcggacgacg ttctcgccac cctggccaag aaggcggaaa aggaggggta cgaggtgcgc 420
atcctcaccg ccgaccgga cctctaccaa ctcgtctccg accgcgtcgc cgtcctccac 480

| | | | | | | |
|------------|------------|------------|-------------|------------|-------------|------|
| cccgagggcc | acctcatcac | cccgagtg | ctttgggaga | agtacggcct | caggccggag | 540 |
| cagtgggtgg | acttccgcgc | cctcgtgggg | gacctctccg | acaacctccc | cgggggtcaag | 600 |
| ggcatcgggg | agaagaccgc | cctcaagctc | ctcaaggagt | ggggaagcct | ggaaaacctc | 660 |
| ctcaagaacc | tggaccgggt | aaagccagaa | aacgtccggg | agaagatcaa | ggcccacctg | 720 |
| gaagacctca | ggctctcctt | ggagctctcc | cgggtgcgca | ccgacctccc | cctggaggtg | 780 |
| gacctcgccc | aggggcggga | gcccgaaccg | gaggggctta | gggccttcct | ggagaggctg | 840 |
| gagttcggca | gcctcctcca | cgagttcggc | ctcctggagg | cccccgcccc | cctggaggag | 900 |
| gccccctggc | ccccgccgga | aggggccttc | gtgggcttcg | tcctctcccc | ccccgagccc | 960 |
| atgtggggcg | agcttaaagc | cctggccgcc | tgcagggacg | gccgggtgca | ccgggcagca | 1020 |
| gacctcttgg | cggggctaaa | ggacctcaag | gaggtccggg | gcctcctcgc | caaggacctc | 1080 |
| gccgtcttgg | cctcgagggg | ggggctagac | ctcgtgcccg | gggacgaccc | catgctcctc | 1140 |
| gcctacctcc | tggacccttc | caacaccacc | cccgaggggg | tggcgcgggc | ctacgggggg | 1200 |
| gagtggacgg | aggacgccgc | ccaccggggc | ctcctctcgg | agaggctcca | tcggaacctc | 1260 |
| cttaagcgcc | tcgaggggga | ggagaagctc | ctttggctct | accacgaggt | ggaaaagccc | 1320 |
| ctctcccggg | tcctggccca | catggaggcc | accggggtag | ggcgggacgt | ggcctacctt | 1380 |
| caggcccttt | ccctggagct | tgcggaggag | atccgccgcc | tcgaggagga | ggtcttccgc | 1440 |
| ttggcggggc | acccttcaa | cctcaactcc | cgggaccagc | tggaaagggg | gctctttgac | 1500 |
| gagcttaggc | ttcccgctt | ggggaagacg | caaaagacag | gcaagcgctc | caccagcgcc | 1560 |
| gcggtgctgg | aggccctacg | ggaggcccac | cccatcgtgg | agaagatcct | ccagcaccgg | 1620 |
| gagctcacca | agctcaagaa | cacctacgtg | gacctctccc | caagcctcgt | ccaccgagg | 1680 |
| acgggcccgc | tcacacccg | cttcaaccag | acggccacgg | ccacggggag | gcttagtagc | 1740 |
| tcgaccccca | acctgcagaa | catccccgtc | cgcacccccct | tgggccagag | gatccgccgg | 1800 |
| gccttcgtgg | ccgaggcggg | ttgggcgttg | gtggccctgg | actatagcca | gatagagctc | 1860 |
| cgcgtcctcg | cccacctctc | cggggacgaa | aacctgatca | gggtcttcca | ggaggggaag | 1920 |
| gacatccaca | cccagaccgc | aagctggatg | ttcggcgtcc | ccccggaggc | cgtggacccc | 1980 |
| ctgatgcgcc | gggcggccaa | gacggtgaac | ttcggcgtcc | tctacggcat | gtccgcccac | 2040 |
| aggctctccc | aggagcttgc | catcccctac | gaggaggcgg | tggcctttat | agagcgctac | 2100 |
| ttccaaagct | tccccaaagg | gcgggcctgg | atagaaaaga | ccctggagga | ggggaggaag | 2160 |

cggggctacg tggaaaccct ctctggaaga aggcgctacg tgcccgcact caacgcccgg 2220
 gtgaagagcg tcagggaggc cgcggagcgc atggccttca acatgcccggt ccagggcacc 2280
 gccgcccacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg 2340
 gcccgcatgc tcctccagggt ccacgacgag ctctctctgg agggccccca agcgcggggc 2400
 gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatcccct cgcggtgccc 2460
 ctggaggtgg aggtggggat gggggaggac tggctttccg ccaagggtta g 2511

<210> 245
 <211> 30
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 245
 atagccatgg tggagcggcc gctctcccgg 30

<210> 246
 <211> 33
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 246
 aagcgtcgac tcaatcctgc ttgcctcca gcc 33

<210> 247
 <211> 32
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 247
 aatcgaattc accccacttt ttgacctgga gg 32

<210> 248
 <211> 21
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 248
ccgggagagc ggccgctcca c 21

<210> 249
<211> 2508
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 249
atggaattca cccactttt tgacctggag gaacccccca agcgggtgct tctggtggac 60
ggccaccacc tggcctaccg caccttctat gccctgagcc tcaccacctc ccggggggag 120
ccggtgcaga tggctctacgg ctctgcccgg agcctcctca aggccttgaa ggaggacgga 180
caggcggtgg tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacgag 240
gcctacaagg cggggccgggc cccacccccg gaggacttcc cccgccagct cgccttggtc 300
aagcggctgg tggaccttct gggcctggtc cgcctcgagg ccccggggta cgaggcggac 360
gacgtcctgg gcacctggc caagaaggcc gaaagggagg ggatggaggt gcgcacctc 420
acgggagacc gggacttctt ccagctctc tccgagaagg tctcggtcct cctgccggac 480
gggaccctgg tcaccccaaa ggacgtccag gagaagtacg ggtgcccccc ggagcgctgg 540
gtggacttcc gcgcctcac gggggaccgc tcggacaaca tccccggggt ggcggggata 600
ggggagaaga ccgcccttcg actcctcgca gagtggggga gcgtggaaaa cctcctgaag 660
aacctggacc gggtaaagcc ggactcgctc cggcgcaaga tagaggcgca cctcgaggac 720
ctccacctct ccttagacct ggcccgcac cgcaccgacc tccccctgga ggtggacttt 780
aaggccctgc gccgcaggac ccccgacctg gagggcctga gggccttttt ggaggagctg 840
gagttcggaa gcctcctcca cgagttcggc ctcttgggag gggagaagcc ccgggaggag 900
gccccctggc ccccgccga aggggccttc gtgggcttcc tcctttcccg caaggagccc 960
atgtgggcgg agcttctggc cctggcggcg gcctcgagg gccgggtcca ccgggcaaca 1020
agcccggttg aggccctggc cgacctcaag gagggccggg ggttcctggc caaggacctg 1080
gccgttttgg ccctgcggga ggggggtggc ctggacccca cggacgacct cctcctggtg 1140
gcctacctcc tggaccggc caacacccac cccgaggggg tggcccggcg ctacgggggc 1200
gagttcacgg aggacgcagc ggagagggcc ctctctctcg agaggctctt ccagaacctc 1260
tttccccggc tttccgagaa gctcctctgg ctctaccagg aggtggagcg gcccctctcc 1320

His Arg Ala Thr Ser Pro Val Glu Ala Leu Ala Asp Leu Lys Glu Ala
 340 345 350
 Arg Gly Phe Leu Ala Lys Asp Leu Ala Val Leu Ala Leu Arg Glu Gly
 355 360 365
 Val Ala Leu Asp Pro Thr Asp Asp Pro Leu Leu Val Ala Tyr Leu Leu
 370 375 380
 Asp Pro Ala Asn Thr His Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly
 385 390 395 400
 Glu Phe Thr Glu Asp Ala Ala Glu Arg Ala Leu Leu Ser Glu Arg Leu
 405 410 415
 Phe Gln Asn Leu Phe Pro Arg Leu Ser Glu Lys Leu Leu Trp Leu Tyr
 420 425 430
 Gln Glu Val Glu Arg Pro Leu Ser Arg Val Leu Ala His Met Glu Ala
 435 440 445
 Arg Gly Val Arg Leu Asp Val Pro Leu Leu Glu Ala Leu Ser Phe Glu
 450 455 460
 Leu Glu Lys Glu Met Glu Arg Leu Glu Gly Glu Val Phe Arg Leu Ala
 465 470 475 480
 Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu
 485 490 495
 Phe Asp Glu Leu Gly Leu Thr Pro Val Gly Arg Thr Glu Lys Thr Gly
 500 505 510
 Lys Arg Ser Thr Ala Gln Gly Ala Leu Glu Ala Leu Arg Gly Ala His
 515 520 525
 Pro Ile Val Glu Leu Ile Leu Gln Tyr Arg Glu Leu Ser Lys Leu Lys
 530 535 540
 Ser Thr Tyr Leu Asp Pro Leu Pro Arg Leu Val His Pro Arg Thr Gly
 545 550 555 560
 Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu
 565 570 575
 Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu
 580 585 590
 Gly Gln Arg Ile Arg Lys Ala Phe Val Ala Glu Glu Gly Trp Leu Leu
 595 600 605
 Leu Ala Ala Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu
 610 615 620
 Ser Gly Asp Glu Asn Leu Lys Arg Val Phe Arg Glu Gly Lys Asp Ile
 625 630 635 640

<223> Synthetic

<400> 252

aacagtcgac ctaggccttg gcggaagcc

30

<210> 253

<211> 2499

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 253

| | |
|--|------|
| atggaattcc tgccctctt tgagccaag ggccgggtgc ttctggtgga cgccaccac | 60 |
| ctggcctacc gtacctttt tgcctgaag ggctcacca ccagccggg ggagccggtc | 120 |
| caggcggtgt acgggtttgc caagagcctt ttgaaggcgc taaggggaaga cggggatgtg | 180 |
| gtgatcgtgg tgtttgacgc caaggcccc tccttcgcc accagacctt cgaggcctac | 240 |
| aaggcggggc gggctccac ccccgaggac tttcccggc agcttgccct tatcaaggag | 300 |
| atggtggacc ttttgggcct ggagcgctc gaggtgccg gctttgaagc ggatgacgtc | 360 |
| ctggctacc tggccaagaa ggcgaaaag gaaggctacg aagtgcgcac cctcaccgcg | 420 |
| gaccgggacc tttaccagct tctttcggag cgaatctcca tccttcaccc ggagggttac | 480 |
| ctgatcacc cggagtggct ttgggagaag tatgggctta agccttcca gtgggtggac | 540 |
| taccgggcct tggccgggga ccttcgcac aacatcccc gcgtgaagg catcggggag | 600 |
| aagacggcgg ccaagctgat ccgggagtgg ggaagcctgg aaaaccttct taagcacctg | 660 |
| gaacaggtga aacctgcctc cgtgcgggag aagatcctta gccacatgga ggacctcaag | 720 |
| ctatccctgg agctatccc ggtgcacacg gacttgctcc ttcagggtgga cttcgcccgg | 780 |
| cgccgggagc cggaccggga ggggcttaag gccttttttg agaggctgga gttcggaagc | 840 |
| ctcctccacg agttcggcct gttggaaagc ccggtggcgg cggaggaagc tccctggccg | 900 |
| cccccgagg gagccttcgt ggggtacgtt ctttcccgc ccgagcccat gtggcgagg | 960 |
| cttaacgcct tggccgcgc ctgggaggga agggtttacc ggcggaggga tcccttgagg | 1020 |
| gccttgccgg ggcttgggga ggtgaggggg cttttggcca aggacctggc ggtgctggcc | 1080 |
| ctgagggaa ggattgccct ggcaccgggc gacgaccca tgctcctgc ctacctctg | 1140 |
| gatccttcca acaccgccc cgaaggggta gcccggcgt acggggggga gtggaccgag | 1200 |
| gaggcggggg aaaggcgct gctttccgaa aggccttacg ccgccctcct ggagcggtt | 1260 |

aagggggagg agaggcttct ttggctttac gaggaggtgg aaaagcccct ttcgcgggtc 1320
 ctggcccaca tggaggccac gggggtacgg ttggatgtgg cctacttaaa ggccctttcc 1380
 ctggaggtgg aggcggagat aaggcgcttc gaggaggagg tccaccgcct ggccgggcat 1440
 cctttcaacc tgaactcccc ggaccagctg gaaaggggtca tctttgacga gcttgggctt 1500
 cccgccatcg gcaagacgga gaagacgggc aagcgctcca ccagcgccgc cgttttgagg 1560
 gccttgcggg aggtcatcc catcgtggac cgcatacttc agtaccggga gctttccaag 1620
 ctcaaggga cctacatcga tcccttgctt gccctgggtcc accccaagac gaaccgcctc 1680
 cacaccgtt tcaaccagac ggccaccgcc acggggaggc ttagcagctc ggatcctaatt 1740
 ctgcaaaaata tccccgtgcg cacccttttg ggccagcgga tccgcggggc cttcgtggcc 1800
 gaggaggggt ggaggctggt ggttttggtac tacagccaga ttgagctcag ggtcctggcg 1860
 cacctttccg gggacgagaa cctaattccg gtcttccagg agggccagga catccacacc 1920
 cagacggcca gctggatggt cggcgtgccc ccagaggccg tggattccct gatgcgccgg 1980
 gcggccaaga ccatcaactt cggcgtcctc tacggcatgt ccgcccaccg gctttcgga 2040
 gagctggcca tccccctacga ggaggcgggt gccttcatcg agcgggtattt ccagagctac 2100
 cccaaggtgc gggcctggat tgagaaaacc ctggcggaag gacgggaacg gggctatgtg 2160
 gaaaccctct ttggccgccc gcgtacgtg cccgacttgg cttcccgggt gaagagcatc 2220
 cgggaggcag cggagcgcat ggccttcaac atgccgtcc aggggaccgc cgcggatttg 2280
 atgaaactgg ccatggtgaa gctctttccc aggcctcagg agctgggggc caggatgctt 2340
 ttgcaggtgc acgacgaact ggtcctcgag gctcccaagg agcaagcggga ggaagtcgcc 2400
 caggaggcca agcggaccat ggaggaggtg tggcccctga aggtgccctt ggaggtggaa 2460
 gtgggcatcg gggaggactg gctttccgcc aaggcctag 2499

<210> 254
 <211> 832
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 254

Met Glu Phe Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu Val
 1 5 10 15

Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu Lys Gly Leu

<400> 256
ggcgggtgcc tggacgggca 20

<210> 257
<211> 20
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 257
ccagctcggt gtggaccta 20

<210> 258
<211> 2505
<212> DNA
<213> Thermus aquaticus

<400> 258
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcgggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggcccgggc cccacgcgcg gaggactttc cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggctggcg cgctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggt caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg 720
aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttgcc 840
agcctcctcc acgagttcgg cttcttgaa agccccaagg ccctggagga ggccccctgg 900
ccccgcgcgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc 960
gatcttctgg ccctggcgcg cgccaggggg ggccgggtcc accgggcccc cgagccttat 1020

aaagccctca gggacctgaa ggaggcgcg gggcttctcg ccaaagacct gagcgttctg 1080
gccctgaggg aaggccttgg cctcccgccc ggcgacgacc ccatgctcct cgctacctc 1140
ctggaccctt ccaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200
gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gtgggggagg 1260
cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct 1320
gtcctggccc acatggaggc cacgggggtg cgctggacg tggcctatct cagggccttg 1380
tccttgaggg tggccgagga gatcgccgcg ctcgaggccg aggtcttccg cctggccggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
cttcccgcca tcggcaagac ggagaagacc ggcaagcgct ccaccagcgc cgccgtcctg 1560
gaggccctcc gcgagggcca ccccatcggt gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgacccttgg ccggacctca tccaccccag gacgggcccgc 1680
ctccacacc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccc ggccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct caggggtgctg 1860
gcccacctct ccggcgacga gaacctgac cggtcttcc aggaggggcg ggacatccac 1920
acggagaccg ccagctggat gttcggcgtc cccggggagg ccgtggaccc cctgatgcgc 1980
cgggcgccca agaccatcaa cttcggggtc ctctacggca tgcgggcca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcggggggtac 2160
gtggagacc tcttcggccg ccgcgcctac gtgccagacc tagaggcccg ggtgaagagc 2220
gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
ctcatgaagc tggctatggt gaagctcttc ccaggctgg aggaaatggg ggccaggatg 2340
ctccttcagg tccacaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
gcccggtgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
gagggtggga taggggagga ctggctctcc gccaaaggat gatag 2505

<210> 259
<211> 833
<212> PRT
<213> Thermus aquaticus

<400> 259

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Asn | Ser | Gly | Met | Leu | Pro | Leu | Phe | Glu | Pro | Lys | Gly | Arg | Val | Leu | 1 | 5 | 10 | 15 |
| Leu | Val | Asp | Gly | His | His | Leu | Ala | Tyr | Arg | Thr | Phe | His | Ala | Leu | Lys | 20 | 25 | 30 | |
| Gly | Leu | Thr | Thr | Ser | Arg | Gly | Glu | Pro | Val | Gln | Ala | Val | Tyr | Gly | Phe | 35 | 40 | 45 | |
| Ala | Lys | Ser | Leu | Leu | Lys | Ala | Leu | Lys | Glu | Asp | Gly | Asp | Ala | Val | Ile | 50 | 55 | 60 | |
| Val | Val | Phe | Asp | Ala | Lys | Ala | Pro | Ser | Phe | Arg | His | Glu | Ala | Tyr | Gly | 65 | 70 | 75 | 80 |
| Gly | Tyr | Lys | Ala | Gly | Arg | Ala | Pro | Thr | Pro | Glu | Asp | Phe | Pro | Arg | Gln | 85 | 90 | 95 | |
| Leu | Ala | Leu | Ile | Lys | Glu | Leu | Val | Asp | Leu | Leu | Gly | Leu | Ala | Arg | Leu | 100 | 105 | 110 | |
| Glu | Val | Pro | Gly | Tyr | Glu | Ala | Asp | Asp | Val | Leu | Ala | Ser | Leu | Ala | Lys | 115 | 120 | 125 | |
| Lys | Ala | Glu | Lys | Glu | Gly | Tyr | Glu | Val | Arg | Ile | Leu | Thr | Ala | Asp | Lys | 130 | 135 | 140 | |
| Asp | Leu | Tyr | Gln | Leu | Leu | Ser | Asp | Arg | Ile | His | Val | Leu | His | Pro | Glu | 145 | 150 | 155 | 160 |
| Gly | Tyr | Leu | Ile | Thr | Pro | Ala | Trp | Leu | Trp | Glu | Lys | Tyr | Gly | Leu | Arg | 165 | 170 | 175 | |
| Pro | Asp | Gln | Trp | Ala | Asp | Tyr | Arg | Ala | Leu | Thr | Gly | Asp | Glu | Ser | Asp | 180 | 185 | 190 | |
| Asn | Leu | Pro | Gly | Val | Lys | Gly | Ile | Gly | Glu | Lys | Thr | Ala | Arg | Lys | Leu | 195 | 200 | 205 | |
| Leu | Glu | Glu | Trp | Gly | Ser | Leu | Glu | Ala | Leu | Leu | Lys | Asn | Leu | Asp | Arg | 210 | 215 | 220 | |
| Leu | Lys | Pro | Ala | Ile | Arg | Glu | Lys | Ile | Leu | Ala | His | Met | Asp | Asp | Leu | 225 | 230 | 235 | 240 |
| Lys | Leu | Ser | Trp | Asp | Leu | Ala | Lys | Val | Arg | Thr | Asp | Leu | Pro | Leu | Glu | 245 | 250 | 255 | |
| Val | Asp | Phe | Ala | Lys | Arg | Arg | Glu | Pro | Asp | Arg | Glu | Arg | Leu | Arg | Ala | 260 | 265 | 270 | |
| Phe | Leu | Glu | Arg | Leu | Glu | Phe | Gly | Ser | Leu | Leu | His | Glu | Phe | Gly | Leu | 275 | 280 | 285 | |
| Leu | Glu | Ser | Pro | Lys | Ala | Leu | Glu | Glu | Ala | Pro | Trp | Pro | Pro | Pro | Glu | 290 | 295 | 300 | |

Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala
305 310 315 320

Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala
325 330 335

Pro Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu
340 345 350

Leu Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu
355 360 365

Pro Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser
370 375 380

Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr
385 390 395 400

Glu Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn
405 410 415

Leu Trp Gly Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg
420 425 430

Glu Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr
435 440 445

Gly Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val
450 455 460

Ala Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly
465 470 475 480

His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe
485 490 495

Asp Glu Leu Gly Leu Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys
500 505 510

Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro
515 520 525

Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser
530 535 540

Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg
545 550 555 560

Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser
565 570 575

Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly
580 585 590

Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val
595 600 605

<210> 261
 <211> 836
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 261

```

Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
1          5          10          15

Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu
20          25          30

Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
35          40          45

Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala
50          55          60

Val Phe Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala
65          70          75          80

Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro
85          90          95

Arg Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr
100         105         110

Arg Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu
115         120         125

Ala Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala
130         135         140

Asp Arg Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His
145         150         155         160

Pro Glu Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly
165         170         175

Leu Arg Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro
180         185         190

Ser Asp Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu
195         200         205

Lys Leu Leu Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu
210         215         220

Asp Arg Val Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu
225         230         235         240

Glu Asp Leu Arg Leu Ser Leu Glu Leu Ser Arg Val Arg Thr Asp Leu

```


545 550 555 560
 Thr Gly Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly
 565 570 575
 Arg Leu Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr
 580 585 590
 Pro Leu Gly Gln Arg Ile Arg Arg Ala Phe Val Ala Glu Ala Gly Trp
 595 600 605
 Ala Leu Val Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala
 610 615 620
 His Leu Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys
 625 630 635 640
 Asp Ile His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu
 645 650 655
 Ala Val Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly
 660 665 670
 Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile
 675 680 685
 Pro Tyr Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe
 690 695 700
 Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys
 705 710 715 720
 Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp
 725 730 735
 Leu Asn Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala
 740 745 750
 Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala
 755 760 765
 Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu
 770 775 780
 Leu Gln Val His Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala
 785 790 795 800
 Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro
 805 810 815
 Leu Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu
 820 825 830
 Ser Ala Lys Gly
 835
 <210> 262

<211> 2511
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 262
 atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
 ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
 gaaccggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac 180
 ggggtacaagg ccgtcttcgt ggtctttgac gccaaaggccc cctccttcgc ccacgaggcc 240
 tacgaggcct acaaggcggg gagggccccg acccccgagg acttcccccg gcagctcgcc 300
 ctcatcaagg agctggtgga cctcctgggg tttaccgcgc tcgaggtccc cggctacgag 360
 gcggaacgacg ttctcgccac cctggccaag aaggcggaaa aggaggggta cgaggtgcgc 420
 atcctcaccg ccgaccgcga cctctaccaa ctgctctccg accgcgtcgc cgtcctccac 480
 cccgagggcc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag 540
 cagtgggtgg acttccgcgc cctcgtgggg gaccctccg acaacctccc cggggtcaag 600
 ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc 660
 ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg 720
 gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctccc cctggagggtg 780
 gacctcgccc aggggcggga gcccgaccgg gaggggctta gggccttcct ggagaggctg 840
 gagttcggca gcctcctcca cgagttcggc ctcttgagg ccccgcccc cctggaggag 900
 gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctccc ccccgagccc 960
 atgtgggcgg agcttaaagc cctggccgcc tgcagggacg gccgggtgca ccgggcagca 1020
 gacccttggt cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc 1080
 gccgtcttgg cctcgaggga ggggctagac ctggtgccc gggacgacct catgctcctc 1140
 gcctacctcc tggaccctc caacaccacc cccgaggggg tggcgcgggc ctacgggggg 1200
 gagtggacgg aggacgccgc ccaccgggccc ctctctcgg agaggctcca tcggaacctc 1260
 cttaagcgcc tcgaggggga ggagaagctc ctttggctct accacgaggt ggaaaagccc 1320
 ctctcccggg tcctggccca catggaggcc accgggggtac ggcgggacgt ggcctacctt 1380
 caggcccttt ccttgagct tgcggaggag atccgccgcc tcgaggagga ggtcttcgc 1440
 ttggcgggcc accccttcaa cctcaactcc cgggaccagc tggaaagggt gctctttgac 1500

tgctgcagg tcgacgctag ctagtgggtgg tgggtgggtgg gacccttggc ggaaagcc 58

<210> 265
 <211> 2517
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 265
 atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
 caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag cgggggggag 120
 ccggtgcagg cgggtctacgg cttecgccaag agcctcctca aggccctcaa ggaggacggg 180
 gacgcggtga tcgtgggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
 ggggtacaagg cggggccgggc cccacgccc gaggactttc cccggcaact cgccctcatc 300
 aaggagctgg tggacctcct ggggctggcg cgctcagagg tcccgggcta cgaggcggac 360
 gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420
 accgccgaca aagaccttta ccagctcctt tccgaccga tccacgtcct ccaccccgag 480
 gggtaacctca tcaccccggc ctggccttg gaaaagtacg gcctgaggcc cgaccagtgg 540
 gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggc caagggcatc 600
 ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
 aacctggacc gggtgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg 720
 aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggc ggacttcgcc 780
 aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
 agcctcctcc acgagtccgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
 ccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggccc 960
 gatcttctgg ccctggccgc cgccaggggg ggcgggtcc accgggcccc cgagccttat 1020
 aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg 1080
 gccctgaggg aaggccttgg cctcccgccc gggacgacc ccatgctcct cgcctacctc 1140
 ctggaccctt ccaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200
 gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gtgggggagg 1260
 cttgaggggg aggagaggct ctttggctt taccgggagg tggagaggcc ctttccgct 1320
 gtcttgccc acatggaggc cacgggggtg cgctggacg tggcctatct cagggccttg 1380

tccctggagg tggccgagga gatcgcccgc ctcgaggccg aggtcttccg cctggccggc 1440
 cacccttca acctcaactc ccgggaccag ctggaaaggg tctcttttga cgagctaggg 1500
 cttcccgcga tcggcaagac ggagaagacc ggcaagcgct ccaccagcgc cgccgtcctg 1560
 gagggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
 aagctgaaga gcacctacat tgaccccttg ccggacctca tccaccccag gacgggcccgc 1680
 ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
 aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccc ggcccttcac 1800
 gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
 gccacacctc ccggcgacga gaacctgatc cgggtcttcc aggagggggc ggacatccac 1920
 acggagaccg ccagctggat gttcggcgtc ccccgaggagg ccgtggaccc cctgatgcgc 1980
 cgggcggcca agaccatcaa cttcggggtc ctctacggca tgcggccca ccgcctctcc 2040
 caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
 gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
 ctcatgaagc tggctatggt gaagctcttc cccaggctgg aggaaatggg ggccaggatg 2340
 ctccctcagg tccacaacga gctggctctc gaggccccaa aagagagggc ggaggccgtg 2400
 gcccggtgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
 gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 266
 <211> 839
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 266

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
 1 5 10 15

Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
 20 25 30

Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe

| 35 | 40 | 45 |
|-------------------------|-------------------------|---------------------|
| Ala Lys Ser Leu Leu Lys | Ala Leu Lys Glu Asp Gly | Asp Ala Val Ile |
| 50 | 55 | 60 |
| Val Val Phe Asp Ala Lys | Ala Pro Ser Phe Arg His | Glu Ala Tyr Gly |
| 65 | 70 | 75 80 |
| Gly Tyr Lys Ala Gly Arg | Ala Pro Thr Pro Glu Asp | Phe Pro Arg Gln |
| 85 | 90 | 95 |
| Leu Ala Leu Ile Lys Glu | Leu Val Asp Leu Leu Gly | Leu Ala Arg Leu |
| 100 | 105 | 110 |
| Glu Val Pro Gly Tyr Glu | Ala Asp Asp Val Leu Ala | Ser Leu Ala Lys |
| 115 | 120 | 125 |
| Lys Ala Glu Lys Glu Gly | Tyr Glu Val Arg Ile Leu | Thr Ala Asp Lys |
| 130 | 135 | 140 |
| Asp Leu Tyr Gln Leu Leu | Ser Asp Arg Ile His Val | Leu His Pro Glu |
| 145 | 150 | 155 160 |
| Gly Tyr Leu Ile Thr Pro | Ala Trp Leu Trp Glu Lys | Tyr Gly Leu Arg |
| 165 | 170 | 175 |
| Pro Asp Gln Trp Ala Asp | Tyr Arg Ala Leu Thr | Gly Asp Glu Ser Asp |
| 180 | 185 | 190 |
| Asn Leu Pro Gly Val Lys | Gly Ile Gly Glu Lys Thr | Ala Arg Lys Leu |
| 195 | 200 | 205 |
| Leu Glu Glu Trp Gly Ser | Leu Glu Ala Leu Leu Lys | Asn Leu Asp Arg |
| 210 | 215 | 220 |
| Leu Lys Pro Ala Ile Arg | Glu Lys Ile Leu Ala His | Met Asp Asp Leu |
| 225 | 230 | 235 240 |
| Lys Leu Ser Trp Asp Leu | Ala Lys Val Arg Thr Asp | Leu Pro Leu Glu |
| 245 | 250 | 255 |
| Val Asp Phe Ala Lys Arg | Arg Glu Pro Asp Arg Glu | Arg Leu Arg Ala |
| 260 | 265 | 270 |
| Phe Leu Glu Arg Leu Glu | Phe Gly Ser Leu Leu His | Glu Phe Gly Leu |
| 275 | 280 | 285 |
| Leu Glu Ser Pro Lys Ala | Leu Glu Glu Ala Pro Trp | Pro Pro Pro Glu |
| 290 | 295 | 300 |
| Gly Ala Phe Val Gly Phe | Val Leu Ser Arg Lys Glu | Pro Met Trp Ala |
| 305 | 310 | 315 320 |
| Asp Leu Leu Ala Leu Ala | Ala Ala Arg Gly Gly Arg | Val His Arg Ala |
| 325 | 330 | 335 |
| Pro Glu Pro Tyr Lys Ala | Leu Arg Asp Leu Lys Glu | Ala Arg Gly Leu |

| | | | | | |
|-------------|---|-----|-----|-----|-----|
| | 645 | | 650 | | 655 |
| Pro Leu Met | Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr | | | | |
| | 660 | | 665 | | 670 |
| Gly Met Ser | Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu | | | | |
| | 675 | | 680 | | 685 |
| Glu Ala Gln | Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val | | | | |
| | 690 | | 695 | | 700 |
| Arg Ala Trp | Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr | | | | |
| | 705 | | 710 | | 715 |
| Val Glu Thr | Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala | | | | |
| | | 725 | | 730 | 735 |
| Arg Val Lys | Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met | | | | |
| | | 740 | | 745 | 750 |
| Pro Val Gln | Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys | | | | |
| | | 755 | | 760 | 765 |
| Leu Phe Pro | Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val | | | | |
| | | 770 | | 775 | 780 |
| His Asn Glu | Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val | | | | |
| | | 785 | | 790 | 795 |
| Ala Arg Leu | Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val | | | | |
| | | 805 | | 810 | 815 |
| Pro Leu Glu | Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys | | | | |
| | | 820 | | 825 | 830 |
| Glu His His | His His His His | | | | |
| | 835 | | | | |

<210> 267
 <211> 2526
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

| | |
|---|-----|
| <400> 267 | |
| atgaattccg aggcgatgct tccgctcttt gaaccctaaag gccgggtcct cctgggtggac | 60 |
| ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccgggggc | 120 |
| gaaccgggtgc aggcgggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac | 180 |
| gggtacaagg ccgtcttcgt ggtctttgac gccaaaggccc cctccttcg ccacgaggcc | 240 |
| tacgaggcct acaaggcggg gagggccccg acccccgagg acttcccccg gcagctcgcc | 300 |

| | |
|--|------|
| ctcatcaagg agctggtgga cctcctgggg tttaccgcgc tcgaggtccc cggtacgag | 360 |
| gcgagcagcg ttctcgccac cctggccaag aaggcggaag aggaggggta cgaggtgcgc | 420 |
| atcctcaccg ccgaccgcga cctctaccaa ctctgtctccg accgcgtcgc cgtcctccac | 480 |
| cccgagggcc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag | 540 |
| cagtgggtgg acttccgcgc cctcgtgggg gacccctccg acaacctccc cggggtcaag | 600 |
| ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc | 660 |
| ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg | 720 |
| gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctccc cctggaggtg | 780 |
| gacctcgcgc aggggcggga gcccgaccgg gaggggctta gggccttctt ggagaggctg | 840 |
| gagttcggca gcctcctcca cgagttcggc ctcttgaggg ccccgcccc cctggaggag | 900 |
| gccccctggc ccccgccgga aggggccttc gtgggcttcg tctctctccg ccccgagccc | 960 |
| atgtgggcgg agcttaaagc cctggccgcc tgcaggagcg gccgggtgca ccgggcagca | 1020 |
| gaccccttgg cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc | 1080 |
| gccgtcttgg cctcgaggga ggggctagac ctctgtcccg gggacgacct catgctctc | 1140 |
| gcctacctcc tggacctctc caacaccacc cccgaggggg tggcgcgcg ctacgggggg | 1200 |
| gagtggacgg aggacgcgc ccaccgggccc ctctctcgg agaggctcca tcggaacctc | 1260 |
| cttaagcgcc tcgaggggga ggagaagctc ctttggctct accacgaggt ggaaaagccc | 1320 |
| ctctcccggg tcttgcccca catggaggcc accggggtag ggcgggacgt ggctacctt | 1380 |
| caggcccttt ccctggagct tgcggaggag atccgccgcc tcgaggagga ggtcttccgc | 1440 |
| ttggcggggc accccttcaa cctcaactcc cgggaccagc tggaaagggt gctctttgac | 1500 |
| gagcttaggc ttccgcctt ggggaagacg caaaagacag gcaagcgctc caccagcgcc | 1560 |
| gcggtgctgg aggccctacg ggaggccac cccatcgtgg agaagatcct ccagcaccgg | 1620 |
| gagctacca agtcaagaa cacctacgtg gacccctcc caagcctcgt ccacccgagg | 1680 |
| acgggcccgc tccacaccg cttcaaccag acggccacgg ccacggggag gcttagtagc | 1740 |
| tccgaccca acctgcagaa catccccgtc cgcacccct tgggcccagag gatccgccgg | 1800 |
| gccttcgtgg ccgaggcggg ttgggcgttg gtggccctgg actatagcca gatagagctc | 1860 |
| cgcgtcctcg ccacctctc cggggacgaa aacctgatca gggctctcca ggaggggaag | 1920 |
| gacatccaca cccagaccgc aagctggatg ttcggcgctc ccccgagggc cgtggacccc | 1980 |
| ctgatgcgcc gggcgggcaa gacggtgaac ttcggcgctc tctacggcat gtccgcccac | 2040 |

Asp Arg Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His
 145 150 155 160
 Pro Glu Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly
 165 170 175
 Leu Arg Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro
 180 185 190
 Ser Asp Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu
 195 200 205
 Lys Leu Leu Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu
 210 215 220
 Asp Arg Val Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu
 225 230 235 240
 Glu Asp Leu Arg Leu Ser Leu Glu Leu Ser Arg Val Arg Thr Asp Leu
 245 250 255
 Pro Leu Glu Val Asp Leu Ala Gln Gly Arg Glu Pro Asp Arg Glu Gly
 260 265 270
 Leu Arg Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu
 275 280 285
 Phe Gly Leu Leu Glu Ala Pro Ala Pro Leu Glu Glu Ala Pro Trp Pro
 290 295 300
 Pro Pro Glu Gly Ala Phe Val Gly Phe Val Leu Ser Arg Pro Glu Pro
 305 310 315 320
 Met Trp Ala Glu Leu Lys Ala Leu Ala Ala Cys Arg Asp Gly Arg Val
 325 330 335
 His Arg Ala Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val
 340 345 350
 Arg Gly Leu Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly
 355 360 365
 Leu Asp Leu Val Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu
 370 375 380
 Asp Pro Ser Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly
 385 390 395 400
 Glu Trp Thr Glu Asp Ala Ala His Arg Ala Leu Leu Ser Glu Arg Leu
 405 410 415
 His Arg Asn Leu Leu Lys Arg Leu Glu Gly Glu Glu Lys Leu Leu Trp
 420 425 430
 Leu Tyr His Glu Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met
 435 440 445

Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala
 755 760 765

Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu
 770 775 780

Leu Gln Val His Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala
 785 790 795 800

Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro
 805 810 815

Leu Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu
 820 825 830

Ser Ala Lys Gly His His His His His His
 835 840

<210> 269
 <211> 30
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 269
 gccgccaggg gcggccgcgt ccaccggggc

30

<210> 270
 <211> 31
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 270
 gcctgcaggg gcggccgcgt gcaccggggc a

31

<210> 271
 <211> 26
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 271
 ctcttgacc cttcgaacac cacc

26

<210> 272
 <211> 23
 <212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 272

gtcctggccc atatggaggc cac

23

<210> 273

<211> 2526

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 273

| | |
|---|------|
| atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac | 60 |
| ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc | 120 |
| gaaccggtgc aggcggteta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac | 180 |
| gggtacaagg ccgtcttcgt ggtctttgac gccaaaggccc cctccttcctg ccacgaggcc | 240 |
| tacgaggcct acaaggcggg gagggccccg acccccgagg acttcccccg gcagctcgcc | 300 |
| ctcatcaagg agctggtgga cctcctgggg tttaccgcgc tcgaggtccc cggctacgag | 360 |
| gcggacgacg ttctcgccac cctggccaag aaggcggaaa aggaggggta cgaggtgcgc | 420 |
| atcctcaccg ccgaccgga cctctacca ctctgtctccg accgctcgc cgtcctccac | 480 |
| cccgagggcc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag | 540 |
| cagtgggtgg acttccgcgc cctcgtgggg gaccctccg acaacctccc cgggggtcaag | 600 |
| ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc | 660 |
| ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg | 720 |
| gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctccc cctggagggtg | 780 |
| gacctcgccc aggggcgga gcccgaccgg gaggggctta gggccttcct ggagaggctg | 840 |
| gagttcggca gcctctcca cgagttcggc ctcttgagg ccccgcccc cctggaggag | 900 |
| gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctcccg ccccgagccc | 960 |
| atgtgggogg agcttaaagc cctggccgcc tgcagggggc gccgcgtgca ccgggcagca | 1020 |
| gacccttgg cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc | 1080 |
| gccgtcttgg cctcgaggga ggggctagac ctctgcccc gggacgacct catgctcctc | 1140 |
| gcctacctcc tggaccttc gaacaccacc cccgaggggg tggcgcgcg ctacgggggg | 1200 |

gagtggacgg aggacgccgc ccaccgggcc ctctctcgg agaggetcca tcggaacctc 1260
 cttaaagcgcc tcgaggggga ggagaagctc ctttggtctt accacgaggt ggaaaagccc 1320
 ctctcccggg tcttgccca tatggaggcc accggggtac ggcgggacgt ggcctacctt 1380
 caggcccttt ccctggagct tgcggaggag atccgccgcc tcgaggagga ggtcttcgc 1440
 ttggcgggcc accccttcaa cctcaactcc cgggaccagc tggaaagggg gctctttgac 1500
 gagcttaggc ttccgcctt gggaagacg caaaagacag gcaagcgctc caccagcgcc 1560
 gcggtgctgg aggccctacg ggaggccac cccatcgtgg agaagatcct ccagcaccgg 1620
 gagctacca agctcaagaa cacctacgtg gacccctcc caagcctcgt ccaccgagg 1680
 acgggccgcc tccacaccg cttcaaccag acggccacgg ccacggggag gcttagtagc 1740
 tccgaccca acctgcagaa catccccgtc cgcacccct tgggccagag gatccgccgg 1800
 gccttcgtgg ccgaggcggg ttgggcgttg gtggccctgg actatagcca gatagagctc 1860
 cgcgtcctcg ccacctctc cggggacgaa aacctgatca gggctctcca ggaggggaag 1920
 gacatccaca ccagaccgc aagctggatg ttcggcgctc ccccgaggc cgtggacccc 1980
 ctgatgcgcc ggcggccaa gacggtgaac ttcggcgctc tctacggcat gtccgcccac 2040
 aggctctccc aggagcttgc catcccctac gaggaggcgg tggcctttat agagcgctac 2100
 ttccaaagct tccccaaggt gcgggccttg atagaaaaga ccctggagga ggggaggaag 2160
 cggggctacg tggaaaccct cttcggaaga aggcgctacg tgcccgaact caacgcccgg 2220
 gtgaagagcg tcagggaggc cgcggagcgc atggcettca acatgcccgt ccagggcacc 2280
 gccgcccacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg 2340
 gcccgcatgc tctccaggt ccacaacgag ctctcctcgg agggcccca agcgcgggcc 2400
 gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatccct cgccgtgccc 2460
 ctggaggtgg aggtggggat gggggaggac tggctttccg ccaagggta ccaccaccac 2520
 caccac 2526

<210> 274
 <211> 842
 <212> PRT
 <213> Artificial
 <220>
 <223> Synthetic
 <400> 274

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Asn | Ser | Glu | Ala | Met | Leu | Pro | Leu | Phe | Glu | Pro | Lys | Gly | Arg | Val | 1 | 5 | 10 | 15 |
| Leu | Leu | Val | Asp | Gly | His | His | Leu | Ala | Tyr | Arg | Thr | Phe | Phe | Ala | Leu | 20 | 25 | 30 | |
| Lys | Gly | Leu | Thr | Thr | Ser | Arg | Gly | Glu | Pro | Val | Gln | Ala | Val | Tyr | Gly | 35 | 40 | 45 | |
| Phe | Ala | Lys | Ser | Leu | Leu | Lys | Ala | Leu | Lys | Glu | Asp | Gly | Tyr | Lys | Ala | 50 | 55 | 60 | |
| Val | Phe | Val | Val | Phe | Asp | Ala | Lys | Ala | Pro | Ser | Phe | Arg | His | Glu | Ala | 65 | 70 | 75 | 80 |
| Tyr | Glu | Ala | Tyr | Lys | Ala | Gly | Arg | Ala | Pro | Thr | Pro | Glu | Asp | Phe | Pro | 85 | 90 | 95 | |
| Arg | Gln | Leu | Ala | Leu | Ile | Lys | Glu | Leu | Val | Asp | Leu | Leu | Gly | Phe | Thr | 100 | 105 | 110 | |
| Arg | Leu | Glu | Val | Pro | Gly | Tyr | Glu | Ala | Asp | Asp | Val | Leu | Ala | Thr | Leu | 115 | 120 | 125 | |
| Ala | Lys | Lys | Ala | Glu | Lys | Glu | Gly | Tyr | Glu | Val | Arg | Ile | Leu | Thr | Ala | 130 | 135 | 140 | |
| Asp | Arg | Asp | Leu | Tyr | Gln | Leu | Val | Ser | Asp | Arg | Val | Ala | Val | Leu | His | 145 | 150 | 155 | 160 |
| Pro | Glu | Gly | His | Leu | Ile | Thr | Pro | Glu | Trp | Leu | Trp | Glu | Lys | Tyr | Gly | 165 | 170 | 175 | |
| Leu | Arg | Pro | Glu | Gln | Trp | Val | Asp | Phe | Arg | Ala | Leu | Val | Gly | Asp | Pro | 180 | 185 | 190 | |
| Ser | Asp | Asn | Leu | Pro | Gly | Val | Lys | Gly | Ile | Gly | Glu | Lys | Thr | Ala | Leu | 195 | 200 | 205 | |
| Lys | Leu | Leu | Lys | Glu | Trp | Gly | Ser | Leu | Glu | Asn | Leu | Leu | Lys | Asn | Leu | 210 | 215 | 220 | |
| Asp | Arg | Val | Lys | Pro | Glu | Asn | Val | Arg | Glu | Lys | Ile | Lys | Ala | His | Leu | 225 | 230 | 235 | 240 |
| Glu | Asp | Leu | Arg | Leu | Ser | Leu | Glu | Leu | Ser | Arg | Val | Arg | Thr | Asp | Leu | 245 | 250 | 255 | |
| Pro | Leu | Glu | Val | Asp | Leu | Ala | Gln | Gly | Arg | Glu | Pro | Asp | Arg | Glu | Gly | 260 | 265 | 270 | |
| Leu | Arg | Ala | Phe | Leu | Glu | Arg | Leu | Glu | Phe | Gly | Ser | Leu | Leu | His | Glu | 275 | 280 | 285 | |
| Phe | Gly | Leu | Leu | Glu | Ala | Pro | Ala | Pro | Leu | Glu | Glu | Ala | Pro | Trp | Pro | 290 | 295 | 300 | |

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Pro | Pro | Glu | Gly | Ala | Phe | Val | Gly | Phe | Val | Leu | Ser | Arg | Pro | Glu | Pro | 305 | 310 | 315 | 320 |
| Met | Trp | Ala | Glu | Leu | Lys | Ala | Leu | Ala | Ala | Cys | Arg | Gly | Gly | Arg | Val | 325 | 330 | 335 | |
| His | Arg | Ala | Ala | Asp | Pro | Leu | Ala | Gly | Leu | Lys | Asp | Leu | Lys | Glu | Val | 340 | 345 | 350 | |
| Arg | Gly | Leu | Leu | Ala | Lys | Asp | Leu | Ala | Val | Leu | Ala | Ser | Arg | Glu | Gly | 355 | 360 | 365 | |
| Leu | Asp | Leu | Val | Pro | Gly | Asp | Asp | Pro | Met | Leu | Leu | Ala | Tyr | Leu | Leu | 370 | 375 | 380 | |
| Asp | Pro | Ser | Asn | Thr | Thr | Pro | Glu | Gly | Val | Ala | Arg | Arg | Tyr | Gly | Gly | 385 | 390 | 395 | 400 |
| Glu | Trp | Thr | Glu | Asp | Ala | Ala | His | Arg | Ala | Leu | Leu | Ser | Glu | Arg | Leu | 405 | 410 | 415 | |
| His | Arg | Asn | Leu | Leu | Lys | Arg | Leu | Glu | Gly | Glu | Glu | Lys | Leu | Leu | Trp | 420 | 425 | 430 | |
| Leu | Tyr | His | Glu | Val | Glu | Lys | Pro | Leu | Ser | Arg | Val | Leu | Ala | His | Met | 435 | 440 | 445 | |
| Glu | Ala | Thr | Gly | Val | Arg | Arg | Asp | Val | Ala | Tyr | Leu | Gln | Ala | Leu | Ser | 450 | 455 | 460 | |
| Leu | Glu | Leu | Ala | Glu | Glu | Ile | Arg | Arg | Leu | Glu | Glu | Glu | Val | Phe | Arg | 465 | 470 | 475 | 480 |
| Leu | Ala | Gly | His | Pro | Phe | Asn | Leu | Asn | Ser | Arg | Asp | Gln | Leu | Glu | Arg | 485 | 490 | 495 | |
| Val | Leu | Phe | Asp | Glu | Leu | Arg | Leu | Pro | Ala | Leu | Gly | Lys | Thr | Gln | Lys | 500 | 505 | 510 | |
| Thr | Gly | Lys | Arg | Ser | Thr | Ser | Ala | Ala | Val | Leu | Glu | Ala | Leu | Arg | Glu | 515 | 520 | 525 | |
| Ala | His | Pro | Ile | Val | Glu | Lys | Ile | Leu | Gln | His | Arg | Glu | Leu | Thr | Lys | 530 | 535 | 540 | |
| Leu | Lys | Asn | Thr | Tyr | Val | Asp | Pro | Leu | Pro | Ser | Leu | Val | His | Pro | Arg | 545 | 550 | 555 | 560 |
| Thr | Gly | Arg | Leu | His | Thr | Arg | Phe | Asn | Gln | Thr | Ala | Thr | Ala | Thr | Gly | 565 | 570 | 575 | |
| Arg | Leu | Ser | Ser | Ser | Asp | Pro | Asn | Leu | Gln | Asn | Ile | Pro | Val | Arg | Thr | 580 | 585 | 590 | |
| Pro | Leu | Gly | Gln | Arg | Ile | Arg | Arg | Ala | Phe | Val | Ala | Glu | Ala | Gly | Trp | 595 | 600 | 605 | |

caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
ccggtgcagg cggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggccgggc cccacgcgc gaggactttc cccggcaact cgcctcatc 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggccacat ggacgatctg 720
aagctctcct gggacctggc caagggtgcg accgacctgc ccctggagggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
agcctcctcc acgagttcgg ctttctggaa agcccaagg ccctggagga ggccccctgg 900
cccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggccc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg 1080
gccctgaggg aaggccttgg cctcccgcgc ggcgacgacc ccatgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200
gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gtggggggagg 1260
cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cttttccgct 1320
gtcctggccc atatggaggc cacgggggtg cgcctggacg tggcctatct cagggccttg 1380
tccttgagg tggccgagga gatcgcccgc ctcgaggccg aggtcttccg cctggccggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
cttcccgcca tcggcaagac ggagaagacc ggcaagcgct ccaccagcgc cgccgtcctg 1560
gaggccctcc gcgaggcca cccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgaccttg ccggacctca tccaccccag gacgggccgc 1680
ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgcgc ggccttcac 1800

gcccaggagg ggtggctatt ggtggccctg gactatagcc agatagagct caggggtgctg 1860
 gcccacctct ccggcgacga gaacctgac cggggtcttcc aggaggggag ggacatccac 1920
 acggagaccg ccagctggat gttcggcgtc ccccgggagg ccgtggaccc cctgatgcgc 1980
 cgggcgggcca agaccatcaa cttcggggtc ctctacggca tgcgggcca ccgcctctcc 2040
 caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 ttccccaagg tgccgggctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
 gtggagaccc tcttcggccg ccgcccgtac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cgcccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
 ctcatgaagc tggctatggt gaagctcttc ccagggtgg aggaaatggg ggccaggatg 2340
 ctcttcagg tccacaacga gctggtcttc gaggcccaa aagagagggc ggaggccgtg 2400
 gcccggtgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggagggtg 2460
 gaggtgggga taggggagga ctggctctcc gccaggagc accaccacca ccaccac 2517

<210> 276
 <211> 839
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 276

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
 1 5 10 15
 Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
 20 25 30
 Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
 35 40 45
 Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
 50 55 60
 Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
 65 70 75 80
 Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
 85 90 95
 Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
 100 105 110

Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
725 730 735

Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
740 745 750

Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
755 760 765

Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
770 775 780

His Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
785 790 795 800

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Glu His His His His His His
835

<210> 277
<211> 27
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 277
gaggagggcgg ggcaccgggc cgccctt

27

<210> 278
<211> 36
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 278
ctttccgaga ggctccatcg gaacctgtgg gggagg

36

<210> 279
<211> 36
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 279

ctcttcgccca acctgcttaa gaggcttgag ggggag

36

<210> 280
<211> 27
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 280
aggccccttt cccgggtcct ggcccat

27

<210> 281
<211> 27
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 281
acgggggtgc gccgggacgt ggcctat

27

<210> 282
<211> 27
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 282
gtggcctatc tccaggcctt gtccctg

27

<210> 283
<211> 27
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 283
ttgtccctgg agcttgccga ggagatc

27

<210> 284
<211> 27
<212> DNA
<213> Artificial

<220>

<223> Synthetic

<400> 284

gccgaggaga tccgccgcct cgaggcc

27

<210> 285

<211> 27

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 285

gcccgcctcg aggaggaggt cttccgc

27

<210> 286

<211> 27

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 286

tttgacgagc taaggcttcc cgccatc

27

<210> 287

<211> 27

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 287

atcgccaaga cgcaaaagac cggcaag

27

<210> 288

<211> 27

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 288

aagatcctgc agcaccggga gctcacc

27

<210> 289

<211> 27

<212> DNA

<213> Artificial
 <220>
 <223> Synthetic
 <400> 289
 accaagctga agaacaccta cattgac 27
 <210> 290
 <211> 27
 <212> DNA
 <213> Artificial
 <220>
 <223> Synthetic
 <400> 290
 aagagcacct acgtggaccc cttgccg 27
 <210> 291
 <211> 39
 <212> DNA
 <213> Artificial
 <220>
 <223> Synthetic
 <400> 291
 attgaccct tgccgagcct cgtccacccc aggacgggc 39
 <210> 292
 <211> 33
 <212> DNA
 <213> Artificial
 <220>
 <223> Synthetic
 <400> 292
 ttcgccaacc tgcttgggag gcttgagggg gag 33
 <210> 293
 <211> 33
 <212> DNA
 <213> Artificial
 <220>
 <223> Synthetic
 <400> 293
 ctcttcgcca acctgtggaa gaggcttgag ggg 33

<210> 294
 <211> 33
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 294
 gcttgcggtc tgggtggcga tgccttccc ctc

33

<210> 295
 <211> 33
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 295
 catgttgaag gccatggcct ccgcggcctc cct

33

<210> 296
 <211> 30
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 296
 caggaggagc tcgttggcga cctggaggag

30

<210> 297
 <211> 45
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 297
 ggagcgcttg cctgtcttct tcgtcttctt caaggcggga ggcct

45

<210> 298
 <211> 33
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 298

gaggaccagc tcgttggcga cctgaaggag cat

33

<210> 299

<211> 33

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 299

gaggggcggg acatcgccac ggagaccgcc agc

33

<210> 300

<211> 33

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 300

cagaacatcc ccgtcgccac cccgcttggg cag

33

<210> 301

<211> 33

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 301

gggcttcccg ccatcaagaa gacggagaag acc

33

<210> 302

<211> 39

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 302

ctagggcttc ccgcatcaa gaagacgcaa aagaccggc

39

<210> 303

<211> 39

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 303
ccggggaaag tcctcctccg tctcggcccg gccgcctt

39

<210> 304
<211> 36
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 304
cgggacctcg aggcgcgtga accccaggag gtccac

36

<210> 305
<211> 18
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 305
ctcctccacg agttcggc

18

<210> 306
<211> 48
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 306
accggtcttc ttcgtcttct tcaacttggg aagcctgagc tcgtcaaa

48

<210> 307
<211> 33
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 307
aagacgaaga agaccggtaa gcgctccacc agc

33

<210> 308
<211> 52
<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 308

gtcgactcta gatcagtggt ggtggtggtg gtgcttggcc gcccggcgca tc 52

<210> 309

<211> 60

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<220>

<221> modified_base

<222> (19)..(42)

<223> The bases in these positions within tH primer are 91% of the base shown and 3% each of the other 3 nucleotides

<400> 309

ggagcgctta ccggtctttt gcgtcttctt gatcttgga agccttagct cgtcaaagag 60

<210> 310

<211> 18

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 310

ctctccacg agttcggc 18

<210> 311

<211> 60

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<220>

<221> modified_base

<222> (19)..(42)

<223> The bases at these positions within tH primer are 91% of the base shown and 3% each of the other 3 nucleotides.

<400> 311

caaaagaccg gtaagcgctc caccagcgcc gccgtcctgg aggccctccg cgaggccac 60

<210> 312
 <211> 52
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 312
 gtcgactcta gatcagtggt ggtggtggtg gtgcttggcc gcccggcgca tc 52

 <210> 313
 <211> 21
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 313
 gtcggagggg tccccacga g 21

 <210> 314
 <211> 17
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 314
 tgtggaattg tgagcgg 17

 <210> 315
 <211> 75
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <220>
 <221> modified_base
 <222> (28)..(59)
 <223> The bases in these positions within tH primer are 91% of the base
 shown and 3% each of the other 3 nucleotides.

 <400> 315
 ctctgtggggg acccctccga caacctcccc ggggtcaagg gcatcgggga gaagaccgcc 60

 ctcaagcttc tcaag 75

<210> 316
<211> 23
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 316
gtggcctcca tatgggccag gac

23

<210> 317
<211> 36
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 317
cgggacctcg aggcgcgtga accccaggag gtccac

36

<210> 318
<211> 39
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 318
ccggggaaag tcctcctccg tctcggcccg gcccgctt

39

<210> 319
<211> 24
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 319
gtcggactcg tcaccggtca gggc

24

<210> 320
<211> 75
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<220>
 <221> modified_base
 <222> (28)..(60)
 <223> The bases in these positions within tH primer are 91% of the base
 shown and 3% each of the other 3 nucleotides.

<400> 320
 ctgaccggtg acgagtccga caaccttccc ggggtcaagg gcatcgggga gaggacggcg 60
 aggaagcttc tggag 75

<210> 321
 <211> 33
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 321
 tagctcctgg gagagggcgt gggccgacat gcc 33

<210> 322
 <211> 34
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 322
 cttccagaac ctcttttaaac ggctttccga gaag 34

<210> 323
 <211> 34
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 323
 cttctcgaa agccgtttaa agaggttctg gaag 34

<210> 324
 <211> 31
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 324 31
ccggtgggcc ggacgcagaa gacgggcaag c

<210> 325
<211> 31
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 325 31
gcttgcccgt cttctgcgtc cggcccaccg g

<210> 326
<211> 31
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 326 31
ctcctccaag tgcacaacga gctgggtctg g

<210> 327
<211> 31
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 327 31
ccaggaccag ctcgttgtgc acttggagga g

<210> 328
<211> 26
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 328 26
gccgcctcc tgaagcggct taaggg

<210> 329
<211> 26
<212> DNA
<213> Artificial

<220>
 <223> Synthetic

<400> 329
 cccttaagcc gcttcaggag ggcggc 26

<210> 330
 <211> 28
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 330
 atcggcaaga cgcagaagac gggcaagc 28

<210> 331
 <211> 28
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 331
 gcttgcccgt cttctgcgtc ttgccgat 28

<210> 332
 <211> 27
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 332
 ttgcagggtgc acaacgaact ggtcctc 27

<210> 333
 <211> 27
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 333
 gaggaccagt tcgttggtgca cctgcaa 27

<210> 334

<211> 36
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 334
 cagaccatga attccacccc actttttgac ctggag 36

 <210> 335
 <211> 36
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 335
 gtggacgcgg ccgcccgcagg ccgcccgcag ggccag 36

 <210> 336
 <211> 36
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 336
 cagaccatga attccctgcc cctctttgag cccaag 36

 <210> 337
 <211> 36
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 337
 gtaaaccgcg ccgccccagg cggcggccaa ggcgtt 36

 <210> 338
 <211> 30
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 338
 gaggtggagc ggcccctctc ccgggttttg 30

<223> Synthetic

<400> 341

Met Glu Phe Thr Pro Leu Phe Asp Leu Glu Glu Pro Pro Lys Arg Val
1 5 10 15

Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Tyr Ala Leu
20 25 30

Ser Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Met Val Tyr Gly Phe
35 40 45

Ala Arg Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Gln Ala Val Val
50 55 60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu
65 70 75 80

Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
85 90 95

Leu Ala Leu Val Lys Arg Leu Val Asp Leu Leu Gly Leu Val Arg Leu
100 105 110

Glu Ala Pro Gly Tyr Glu Ala Asp Asp Val Leu Gly Thr Leu Ala Lys
115 120 125

Lys Ala Glu Arg Glu Gly Met Glu Val Arg Ile Leu Thr Gly Asp Arg
130 135 140

Asp Phe Phe Gln Leu Leu Ser Glu Lys Val Ser Val Leu Leu Pro Asp
145 150 155 160

Gly Thr Leu Val Thr Pro Lys Asp Val Gln Glu Lys Tyr Gly Val Pro
165 170 175

Pro Glu Arg Trp Val Asp Phe Arg Ala Leu Thr Gly Asp Arg Ser Asp
180 185 190

Asn Ile Pro Gly Val Ala Gly Ile Gly Glu Lys Thr Ala Leu Arg Leu
195 200 205

Leu Ala Glu Trp Gly Ser Val Glu Asn Leu Leu Lys Asn Leu Asp Arg
210 215 220

Val Lys Pro Asp Ser Leu Arg Arg Lys Ile Glu Ala His Leu Glu Asp
225 230 235 240

Leu His Leu Ser Leu Asp Leu Ala Arg Ile Arg Thr Asp Leu Pro Leu
245 250 255

Glu Val Asp Phe Lys Ala Leu Arg Arg Arg Thr Pro Asp Leu Glu Gly
260 265 270

Leu Arg Ala Phe Leu Glu Glu Leu Glu Phe Gly Ser Leu Leu His Glu
275 280 285

| | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Phe | Gly | Leu | Leu | Gly | Gly | Glu | Lys | Pro | Arg | Glu | Glu | Ala | Pro | Trp | Pro | 290 | 295 | 300 |
| Pro | Pro | Glu | Gly | Ala | Phe | Val | Gly | Phe | Leu | Leu | Ser | Arg | Lys | Glu | Pro | 305 | 310 | 315 |
| Met | Trp | Ala | Glu | Leu | Leu | Ala | Leu | Ala | Ala | Ala | Ser | Glu | Gly | Arg | Val | 325 | 330 | 335 |
| His | Arg | Ala | Thr | Ser | Pro | Val | Glu | Ala | Leu | Ala | Asp | Leu | Lys | Glu | Ala | 340 | 345 | 350 |
| Arg | Gly | Phe | Leu | Ala | Lys | Asp | Leu | Ala | Val | Leu | Ala | Leu | Arg | Glu | Gly | 355 | 360 | 365 |
| Val | Ala | Leu | Asp | Pro | Thr | Asp | Asp | Pro | Leu | Leu | Val | Ala | Tyr | Leu | Leu | 370 | 375 | 380 |
| Asp | Pro | Ala | Asn | Thr | His | Pro | Glu | Gly | Val | Ala | Arg | Arg | Tyr | Gly | Gly | 385 | 390 | 395 |
| Glu | Phe | Thr | Glu | Asp | Ala | Ala | Glu | Arg | Ala | Leu | Leu | Ser | Glu | Arg | Leu | 405 | 410 | 415 |
| Phe | Gln | Asn | Leu | Phe | Lys | Arg | Leu | Ser | Glu | Lys | Leu | Leu | Trp | Leu | Tyr | 420 | 425 | 430 |
| Gln | Glu | Val | Glu | Arg | Pro | Leu | Ser | Arg | Val | Leu | Ala | His | Met | Glu | Ala | 435 | 440 | 445 |
| Arg | Gly | Val | Arg | Leu | Asp | Val | Pro | Leu | Leu | Glu | Ala | Leu | Ser | Phe | Glu | 450 | 455 | 460 |
| Leu | Glu | Lys | Glu | Met | Glu | Arg | Leu | Glu | Gly | Glu | Val | Phe | Arg | Leu | Ala | 465 | 470 | 475 |
| Gly | His | Pro | Phe | Asn | Leu | Asn | Ser | Arg | Asp | Gln | Leu | Glu | Arg | Val | Leu | 485 | 490 | 495 |
| Phe | Asp | Glu | Leu | Gly | Leu | Thr | Pro | Val | Gly | Arg | Thr | Gln | Lys | Thr | Gly | 500 | 505 | 510 |
| Lys | Arg | Ser | Thr | Ala | Gln | Gly | Ala | Leu | Glu | Ala | Leu | Arg | Gly | Ala | His | 515 | 520 | 525 |
| Pro | Ile | Val | Glu | Leu | Ile | Leu | Gln | Tyr | Arg | Glu | Leu | Ser | Lys | Leu | Lys | 530 | 535 | 540 |
| Ser | Thr | Tyr | Leu | Asp | Pro | Leu | Pro | Arg | Leu | Val | His | Pro | Arg | Thr | Gly | 545 | 550 | 555 |
| Arg | Leu | His | Thr | Arg | Phe | Asn | Gln | Thr | Ala | Thr | Ala | Thr | Gly | Arg | Leu | 565 | 570 | 575 |
| Ser | Ser | Ser | Asp | Pro | Asn | Leu | Gln | Asn | Ile | Pro | Val | Arg | Thr | Pro | Leu | 580 | 585 | 590 |

Gly Gln Arg Ile Arg Lys Ala Phe Val Ala Glu Glu Gly Trp Leu Leu
595 600 605

Leu Ala Ala Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu
610 615 620

Ser Gly Asp Glu Asn Leu Lys Arg Val Phe Arg Glu Gly Lys Asp Ile
625 630 635 640

His Thr Glu Thr Ala Ala Trp Met Phe Gly Leu Asp Pro Ala Leu Val
645 650 655

Asp Pro Lys Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly Val Leu
660 665 670

Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Gly Ile Asp Tyr
675 680 685

Lys Glu Ala Glu Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys
690 695 700

Val Arg Ala Trp Ile Glu Arg Thr Leu Glu Glu Gly Arg Thr Arg Gly
705 710 715 720

Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Ala
725 730 735

Ser Arg Val Arg Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn
740 745 750

Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Ile Ala Met Val
755 760 765

Lys Leu Phe Pro Arg Leu Lys Pro Leu Gly Ala His Leu Leu Leu Gln
770 775 780

Val His Asn Glu Leu Val Leu Glu Val Pro Glu Asp Arg Ala Glu Glu
785 790 795 800

Ala Lys Ala Leu Val Lys Glu Val Met Glu Asn Ala Tyr Pro Leu Asp
805 810 815

Val Pro Leu Glu Val Glu Val Gly Val Gly Arg Asp Trp Leu Glu Ala
820 825 830

Lys Gln Asp
835

<210> 342
<211> 38
<212> DNA
<213> Artificial

<220>
<223> Synthetic

| | | | | | | |
|------------|------------|------------|------------|------------|------------|------|
| ggggagaaga | ccgcccttcg | actcctcgca | gagtggggga | gcgtggaaaa | cctcctgaag | 660 |
| aacctggacc | gggtaaagcc | ggactcgctc | cggcgcaaga | tagaggcgca | cctcgaggac | 720 |
| ctccacctct | ccttagacct | ggcccgcatc | cgcaccgacc | ccccctgga | ggtggacttt | 780 |
| aaggccctgc | gccgcaggac | ccccgacctg | gagggcctga | gggccttttt | ggaggagctg | 840 |
| gagttcggaa | gcctcctcca | cgagttcggc | ctcctgggag | gggagaagcc | ccgggaggag | 900 |
| gccccctggc | ccccgcccga | aggggccttc | gtgggcttcc | tcctttcccg | caaggagccc | 960 |
| atgtgggcgg | agcttctggc | cctggcgggc | gcctcggggc | gccgcgtcca | ccgggcaaca | 1020 |
| agcccggttg | aggccctggc | cgacctcaag | gaggcccggg | ggttcctggc | caaggacctg | 1080 |
| gccgttttgg | ccctgcggga | gggggtggcc | ctggacccca | cggacgaccc | cctcctggtg | 1140 |
| gcctacctcc | tggacccggc | caacacccac | cccagggggg | tggcccggcg | ctacgggggc | 1200 |
| gagttcacgg | aggacgcagc | ggagagggcc | ctcctctccg | agaggctctt | ccagaacctc | 1260 |
| tttaaacggc | tttccgagaa | gctcctctgg | ctctaccagg | aggtggagcg | gccccctctc | 1320 |
| cgggtcttgg | cccacatgga | ggcccggggg | gtgaggctgg | acgtccccct | tctggaggcc | 1380 |
| ctctcctttg | agctggagaa | ggagatggag | cgcctggagg | gggaggctct | ccgtttggcc | 1440 |
| ggccaccctt | tcaacctcaa | ctcccgcgac | cagctggaaa | gggtcctctt | tgacgagctg | 1500 |
| ggcctcacc | cggtagggcc | gacgcagaag | acgggcaagc | gctccaccgc | ccagggggcc | 1560 |
| ctggaggccc | tccggggggc | ccaccccatc | gtggagctca | tcctccagta | ccgggagctt | 1620 |
| tccaagctca | aaagcaccta | cctggacccc | ctgccccggc | tcgtccaccc | gcggacgggc | 1680 |
| cggctccaca | cccgttcaa | ccagacggcc | acggccacgg | gaaggctttc | cagctccgac | 1740 |
| cccaacctgc | agaacatccc | cgtgcgcacc | cccttggggc | agcgcatccg | caaggccttc | 1800 |
| gtggccgagg | aggggtggct | ccttttggcg | gcggactact | cccagattga | gctccgggtc | 1860 |
| ctggcccacc | tctcggggga | cgagaacctg | aagcgggtct | tccgggaggg | gaaggacatc | 1920 |
| cataccgaga | ccgccgcctg | gatgttcggc | ttagaccccg | ctctggtgga | tccaaagatg | 1980 |
| cgccggggcg | ccaagacggt | caacttcggc | gtcctctacg | ggatgtccgc | ccacaggctc | 2040 |
| tcccaggagc | tcggcataga | ctacaaggag | gcggaggcct | ttattgagcg | ctacttccag | 2100 |
| agcttcccca | aggtgcgggc | ctggatagaa | aggaccctgg | aggagggccg | gacgcggggc | 2160 |
| tacgtggaga | ccctgttcgg | caggaggcgc | tatgtgcccg | acctggcctc | ccgggtccgc | 2220 |
| tcggtgcggg | aggcggcgga | gcggatggcc | ttcaacatgc | ccgtgcaggg | caccgccgcc | 2280 |
| gacctgatga | agatcgccat | ggtcaagctc | ttccccaggc | taaagcccct | gggggcccac | 2340 |

ctcctcctcc aagtgcacaa cgagctgggtc ctggaggtgc ccgaggaccg ggccgaggag 2400
 gccaaagggccc tgggtcaagga ggtcatggag aacgcctacc ccctggacgt gcccctcgag 2460
 gtggaggtgg gcgtgggtcg ggactggctg gaggcgaagc aggat 2505

<210> 346
 <211> 835
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 346

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Asn | Ser | Thr | Pro | Leu | Phe | Asp | Leu | Glu | Glu | Pro | Pro | Lys | Arg | Val |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Leu | Leu | Val | Asp | Gly | His | His | Leu | Ala | Tyr | Arg | Thr | Phe | Tyr | Ala | Leu |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Ser | Leu | Thr | Thr | Ser | Arg | Gly | Glu | Pro | Val | Gln | Met | Val | Tyr | Gly | Phe |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Ala | Arg | Ser | Leu | Leu | Lys | Ala | Leu | Lys | Glu | Asp | Gly | Gln | Ala | Val | Val |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| Val | Val | Phe | Asp | Ala | Lys | Ala | Pro | Ser | Phe | Arg | His | Glu | Ala | Tyr | Glu |
| 65 | | | | | 70 | | | | | 75 | | | | | 80 |
| Ala | Tyr | Lys | Ala | Gly | Arg | Ala | Pro | Thr | Pro | Glu | Asp | Phe | Pro | Arg | Gln |
| | | | | 85 | | | | | 90 | | | | | 95 | |
| Leu | Ala | Leu | Val | Lys | Arg | Leu | Val | Asp | Leu | Leu | Gly | Leu | Val | Arg | Leu |
| | | | 100 | | | | | 105 | | | | | 110 | | |
| Glu | Ala | Pro | Gly | Tyr | Glu | Ala | Asp | Asp | Val | Leu | Gly | Thr | Leu | Ala | Lys |
| | | 115 | | | | | 120 | | | | | 125 | | | |
| Lys | Ala | Glu | Arg | Glu | Gly | Met | Glu | Val | Arg | Ile | Leu | Thr | Gly | Asp | Arg |
| | | 130 | | | | 135 | | | | | 140 | | | | |
| Asp | Phe | Phe | Gln | Leu | Leu | Ser | Glu | Lys | Val | Ser | Val | Leu | Leu | Pro | Asp |
| 145 | | | | | 150 | | | | | 155 | | | | | 160 |
| Gly | Thr | Leu | Val | Thr | Pro | Lys | Asp | Val | Gln | Glu | Lys | Tyr | Gly | Val | Pro |
| | | | | 165 | | | | | 170 | | | | | 175 | |
| Pro | Glu | Arg | Trp | Val | Asp | Phe | Arg | Ala | Leu | Thr | Gly | Asp | Arg | Ser | Asp |
| | | | 180 | | | | | 185 | | | | | 190 | | |
| Asn | Ile | Pro | Gly | Val | Ala | Gly | Ile | Gly | Glu | Lys | Thr | Ala | Leu | Arg | Leu |
| | | 195 | | | | | 200 | | | | | 205 | | | |

Leu Ala Glu Trp Gly Ser Val Glu Asn Leu Leu Lys Asn Leu Asp Arg
 210 215 220
 Val Lys Pro Asp Ser Leu Arg Arg Lys Ile Glu Ala His Leu Glu Asp
 225 230 235 240
 Leu His Leu Ser Leu Asp Leu Ala Arg Ile Arg Thr Asp Leu Pro Leu
 245 250 255
 Glu Val Asp Phe Lys Ala Leu Arg Arg Arg Thr Pro Asp Leu Glu Gly
 260 265 270
 Leu Arg Ala Phe Leu Glu Glu Leu Glu Phe Gly Ser Leu Leu His Glu
 275 280 285
 Phe Gly Leu Leu Gly Gly Glu Lys Pro Arg Glu Glu Ala Pro Trp Pro
 290 295 300
 Pro Pro Glu Gly Ala Phe Val Gly Phe Leu Leu Ser Arg Lys Glu Pro
 305 310 315 320
 Met Trp Ala Glu Leu Leu Ala Leu Ala Ala Ala Ser Gly Gly Arg Val
 325 330 335
 His Arg Ala Thr Ser Pro Val Glu Ala Leu Ala Asp Leu Lys Glu Ala
 340 345 350
 Arg Gly Phe Leu Ala Lys Asp Leu Ala Val Leu Ala Leu Arg Glu Gly
 355 360 365
 Val Ala Leu Asp Pro Thr Asp Asp Pro Leu Leu Val Ala Tyr Leu Leu
 370 375 380
 Asp Pro Ala Asn Thr His Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly
 385 390 395 400
 Glu Phe Thr Glu Asp Ala Ala Glu Arg Ala Leu Leu Ser Glu Arg Leu
 405 410 415
 Phe Gln Asn Leu Phe Lys Arg Leu Ser Glu Lys Leu Leu Trp Leu Tyr
 420 425 430
 Gln Glu Val Glu Arg Pro Leu Ser Arg Val Leu Ala His Met Glu Ala
 435 440 445
 Arg Gly Val Arg Leu Asp Val Pro Leu Leu Glu Ala Leu Ser Phe Glu
 450 455 460
 Leu Glu Lys Glu Met Glu Arg Leu Glu Gly Glu Val Phe Arg Leu Ala
 465 470 475 480
 Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu
 485 490 495
 Phe Asp Glu Leu Gly Leu Thr Pro Val Gly Arg Thr Gln Lys Thr Gly
 500 505 510

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Lys | Arg | Ser | Thr | Ala | Gln | Gly | Ala | Leu | Glu | Ala | Leu | Arg | Gly | Ala | His | 515 | 520 | 525 | |
| Pro | Ile | Val | Glu | Leu | Ile | Leu | Gln | Tyr | Arg | Glu | Leu | Ser | Lys | Leu | Lys | 530 | 535 | 540 | |
| Ser | Thr | Tyr | Leu | Asp | Pro | Leu | Pro | Arg | Leu | Val | His | Pro | Arg | Thr | Gly | 545 | 550 | 555 | 560 |
| Arg | Leu | His | Thr | Arg | Phe | Asn | Gln | Thr | Ala | Thr | Ala | Thr | Gly | Arg | Leu | 565 | 570 | 575 | |
| Ser | Ser | Ser | Asp | Pro | Asn | Leu | Gln | Asn | Ile | Pro | Val | Arg | Thr | Pro | Leu | 580 | 585 | 590 | |
| Gly | Gln | Arg | Ile | Arg | Lys | Ala | Phe | Val | Ala | Glu | Glu | Gly | Trp | Leu | Leu | 595 | 600 | 605 | |
| Leu | Ala | Ala | Asp | Tyr | Ser | Gln | Ile | Glu | Leu | Arg | Val | Leu | Ala | His | Leu | 610 | 615 | 620 | |
| Ser | Gly | Asp | Glu | Asn | Leu | Lys | Arg | Val | Phe | Arg | Glu | Gly | Lys | Asp | Ile | 625 | 630 | 635 | 640 |
| His | Thr | Glu | Thr | Ala | Ala | Trp | Met | Phe | Gly | Leu | Asp | Pro | Ala | Leu | Val | 645 | 650 | 655 | |
| Asp | Pro | Lys | Met | Arg | Arg | Ala | Ala | Lys | Thr | Val | Asn | Phe | Gly | Val | Leu | 660 | 665 | 670 | |
| Tyr | Gly | Met | Ser | Ala | His | Arg | Leu | Ser | Gln | Glu | Leu | Gly | Ile | Asp | Tyr | 675 | 680 | 685 | |
| Lys | Glu | Ala | Glu | Ala | Phe | Ile | Glu | Arg | Tyr | Phe | Gln | Ser | Phe | Pro | Lys | 690 | 695 | 700 | |
| Val | Arg | Ala | Trp | Ile | Glu | Arg | Thr | Leu | Glu | Glu | Gly | Arg | Thr | Arg | Gly | 705 | 710 | 715 | 720 |
| Tyr | Val | Glu | Thr | Leu | Phe | Gly | Arg | Arg | Arg | Tyr | Val | Pro | Asp | Leu | Ala | 725 | 730 | 735 | |
| Ser | Arg | Val | Arg | Ser | Val | Arg | Glu | Ala | Ala | Glu | Arg | Met | Ala | Phe | Asn | 740 | 745 | 750 | |
| Met | Pro | Val | Gln | Gly | Thr | Ala | Ala | Asp | Leu | Met | Lys | Ile | Ala | Met | Val | 755 | 760 | 765 | |
| Lys | Leu | Phe | Pro | Arg | Leu | Lys | Pro | Leu | Gly | Ala | His | Leu | Leu | Leu | Gln | 770 | 775 | 780 | |
| Val | His | Asn | Glu | Leu | Val | Leu | Glu | Val | Pro | Glu | Asp | Arg | Ala | Glu | Glu | 785 | 790 | 795 | 800 |
| Ala | Lys | Ala | Leu | Val | Lys | Glu | Val | Met | Glu | Asn | Ala | Tyr | Pro | Leu | Asp | 805 | 810 | 815 | |

Val Pro Leu Glu Val Glu Val Gly Val Gly Arg Asp Trp Leu Glu Ala
820 825 830

Lys Gln Asp
835

<210> 347
<211> 2496
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 347
atgaattccc tgccccctctt tgagcccaag ggccgggtgc ttctggtgga cggccaccac 60
ctggcctacc gtaccttttt tgccctgaag ggcctcacca ccagccgcgg ggagccggtc 120
caggcgggtgt acgggttttg caagagcctt ttgaaggcgc taagggaaga cggggatgtg 180
gtgatcgtgg tgtttgacgc caaggccccc tccttcgcc accagaccta cgaggcctac 240
aaggcggggc gggctccac ccccgaggac tttcccggc agcttgccct tatcaaggag 300
atggtggacc ttttgggcct ggagcgcctc gaggtgcgg gctttgaagc ggatgacgtc 360
ctggctaccc tggccaagaa ggcggaaaag gaaggctacg aagtgcgcat cctcaccgcg 420
gaccgggacc tttaccagct tctttcggag cgaatctcca tccttcaccc ggagggttac 480
ctgatcaccc cggagtggct ttgggagaag tatgggctta agccttccca gtgggtggac 540
taccgggcct tggccgggga cccttcgac aacatccccg gcgtgaaggg catcggggag 600
aagacggcgg ccaagctgat ccgggagtgg ggaagcctgg aaaaccttct taagcacctg 660
gaacaggtga aacctgcctc cgtgcgggag aagatcctta gccacatgga ggacctcaag 720
ctatccctgg agctatcccc ggtgcacacg gacttgctcc ttcaggtgga cttcgccccg 780
cgccgggagc cggaccggga ggggcttaag gccttttttg agaggctgga gttcgggaagc 840
ctcctccacg agttcggcct gttggaaagc ccggtggcgg cggaggaagc tccctggccg 900
cccccgagg gagccttcgt ggggtacgtt ctttcccgcc ccgagcccat gtgggcggag 960
cttaacgcct tggccgcgc ctggggcggc cgcgtttacc gggcggagga tcccttggag 1020
gccttgccgg ggcttgggga ggtgaggggg cttttggcca aggacctggc ggtgctggcc 1080
ctgagggaaag ggattgccct ggcaccgggc gacgacccca tgctcctcgc ctacctcctg 1140
gatccttcca acaccgcccc cgaaggggta gcccgcgct acggggggga gtggaccgag 1200
gaggcggggg aaaggcgct gctttccgaa aggctttacg ccgccctcct gaagcggctt 1260

| | | | | | | | | | | | | | | | | | | | |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|--|--|--|--|
| | | | 20 | | | | | | 25 | | | | | | 30 | | | | |
| Thr | Thr | Ser 35 | Arg | Gly | Glu | Pro | Val 40 | Gln | Ala | Val | Tyr | Gly 45 | Phe | Ala | Lys | | | | |
| Ser | Leu 50 | Leu | Lys | Ala | Leu | Arg 55 | Glu | Asp | Gly | Asp | Val 60 | Val | Ile | Val | Val | | | | |
| Phe 65 | Asp | Ala | Lys | Ala | Pro 70 | Ser | Phe | Arg | His | Gln 75 | Thr | Tyr | Glu | Ala | Tyr 80 | | | | |
| Lys | Ala | Gly | Arg | Ala 85 | Pro | Thr | Pro | Glu | Asp 90 | Phe | Pro | Arg | Gln | Leu 95 | Ala | | | | |
| Leu | Ile | Lys | Glu 100 | Met | Val | Asp | Leu | Leu 105 | Gly | Leu | Glu | Arg | Leu 110 | Glu | Val | | | | |
| Pro | Gly | Phe 115 | Glu | Ala | Asp | Asp | Val 120 | Leu | Ala | Thr | Leu | Ala 125 | Lys | Lys | Ala | | | | |
| Glu | Lys 130 | Glu | Gly | Tyr | Glu | Val 135 | Arg | Ile | Leu | Thr | Ala 140 | Asp | Arg | Asp | Leu | | | | |
| Tyr 145 | Gln | Leu | Leu | Ser 150 | Glu | Arg | Ile | Ser | Ile | Leu 155 | His | Pro | Glu | Gly 160 | Tyr | | | | |
| Leu | Ile | Thr | Pro 165 | Glu | Trp | Leu | Trp | Glu | Lys 170 | Tyr | Gly | Leu | Lys | Pro 175 | Ser | | | | |
| Gln | Trp | Val 180 | Asp | Tyr | Arg | Ala | Leu 185 | Ala | Gly | Asp | Pro | Ser | Asp 190 | Asn | Ile | | | | |
| Pro | Gly 195 | Val | Lys | Gly | Ile | Gly | Glu 200 | Lys | Thr | Ala | Ala 205 | Lys | Leu | Ile | Arg | | | | |
| Glu | Trp 210 | Gly | Ser | Leu | Glu | Asn 215 | Leu | Leu | Lys | His 220 | Leu | Glu | Gln | Val | Lys | | | | |
| Pro 225 | Ala | Ser | Val | Arg | Glu 230 | Lys | Ile | Leu | Ser | His 235 | Met | Glu | Asp | Leu 240 | Lys | | | | |
| Leu | Ser | Leu | Glu 245 | Leu | Ser | Arg | Val | His | Thr 250 | Asp | Leu | Leu | Leu | Gln 255 | Val | | | | |
| Asp | Phe | Ala 260 | Arg | Arg | Arg | Glu | Pro 265 | Asp | Arg | Glu | Gly | Leu 270 | Lys | Ala | Phe | | | | |
| Leu | Glu 275 | Arg | Leu | Glu | Phe | Gly | Ser 280 | Leu | Leu | His | Glu | Phe 285 | Gly | Leu | Leu | | | | |
| Glu | Ser 290 | Pro | Val | Ala | Ala | Glu 295 | Glu | Ala | Pro | Trp | Pro 300 | Pro | Pro | Glu | Gly | | | | |
| Ala 305 | Phe | Val | Gly | Tyr 310 | Val | Leu | Ser | Arg | Pro | Glu 315 | Pro | Met | Trp | Ala | Glu 320 | | | | |
| Leu | Asn | Ala | Leu | Ala | Ala | Ala | Trp | Gly | Gly | Arg | Val | Tyr | Arg | Ala | Glu | | | | |

| | | | | | | |
|---|-----|-----|-----|-----|--|-----|
| 625 | | 630 | | 635 | | 640 |
| Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu Ala Val Asp Ser | | | | | | |
| | 645 | | | 650 | | 655 |
| Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr Gly | | | | | | |
| | 660 | | 665 | | | 670 |
| Met Ser Ala His Arg Leu Ser Gly Glu Leu Ala Ile Pro Tyr Glu Glu | | | | | | |
| | 675 | | 680 | | | 685 |
| Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Tyr Pro Lys Val Arg | | | | | | |
| | 690 | | 695 | | | 700 |
| Ala Trp Ile Glu Lys Thr Leu Ala Glu Gly Arg Glu Arg Gly Tyr Val | | | | | | |
| 705 | | 710 | | 715 | | 720 |
| Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Ala Ser Arg | | | | | | |
| | 725 | | 730 | | | 735 |
| Val Lys Ser Ile Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro | | | | | | |
| | 740 | | 745 | | | 750 |
| Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu | | | | | | |
| | 755 | | 760 | | | 765 |
| Phe Pro Arg Leu Gln Glu Leu Gly Ala Arg Met Leu Leu Gln Val His | | | | | | |
| | 770 | | 775 | | | 780 |
| Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Gln Ala Glu Glu Val Ala | | | | | | |
| 785 | | 790 | | 795 | | 800 |
| Gln Glu Ala Lys Arg Thr Met Glu Glu Val Trp Pro Leu Lys Val Pro | | | | | | |
| | 805 | | 810 | | | 815 |
| Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys Ala | | | | | | |
| | 820 | | 825 | | | 830 |

<210> 349
 <211> 2526
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

| | |
|---|-----|
| <400> 349 | |
| atgaattcca ccccaactttt tgacctggag gaacccccca agcgggtgct tctggtggac | 60 |
| ggccaccacc tggcctaccg caccttctat gcctgagcc tcaccacctc ccgggggggag | 120 |
| ccggtgcaga tgggtctacgg cttcgcccgg agcctcctca aggccttgaa ggaggacgga | 180 |
| caggcgggtgg tcgtggtctt tgacgccaag gccccctcct tccgccacga ggccctacgag | 240 |
| gcctacaagg cggggccgggc cccacccccg gaggacttcc cccgccagct cgccttggtc | 300 |

aggctctccc aggagcttgc catccctac gaggaggcgg tggcctttat agagegctac 2100
 ttccaaagct tccccaaggt gcgggcctgg atagaaaaga ccctggagga ggggaggaag 2160
 cggggctacg tggaaaccct cttcggaaga aggcgctacg tgcccgcact caacgcccgg 2220
 gtgaagagcg tcagggaggc cgcggagcgc atggccttca acatgcccgt ccagggcacc 2280
 gccgcccacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg 2340
 gcccgcatgc tcctccaggt cgccaacgag ctctcctgg agggccccca agcgcggggc 2400
 gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatccctc cgccgtgccc 2460
 ctggaggtgg aggtggggat gggggaggac tggctttccg ccaagggtca ccaccaccac 2520
 caccac 2526

<210> 350
 <211> 2505
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 350
 atgaattcca cccactttt tgacctggag gaacccccca agcgggtgct tctggtggac 60
 gggcaccacc tggcctaccg caccttctat gccctgagcc tcaccacctc cgggggggag 120
 ccggtgcaga tggcttacgg cttcgcccgg agcctcctca aggccttgaa ggaggacgga 180
 caggcggtgg tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacgag 240
 gcctacaagg cgggcccggc cccacccccg gaggacttcc cccgccagct cgccttggtc 300
 aagcggctgg tggaccttct gggttttacc cgcctcgagg ccccggggta cgaggcggac 360
 gacgtcctgg gcaccctggc caagaaggcc gaaagggagg ggatggaggt gcgcatcctc 420
 acgggagacc gggacttctt ccagctcctc tccgagaagg tctcggtcct cctgccggac 480
 gggaccctgg tcaccccaaa ggacgtccag gagaagtacg gggtgcccc ggagcgctgg 540
 gtggacttcc gcgccctcac gggggaccgc tcggacaaca tccccggggg ggcggggata 600
 ggggagaaga ccgcccttcg actcctcgca gagtggggga gcgtggaaaa cctcctgaag 660
 aacctggacc gggtaaagcc ggactcgctc cggcgcaaga tagaggcgca cctcgaggac 720
 ctccacctct ccttagacct ggcccgcac cgcaccgacc tccccctgga ggtggacttt 780
 aaggccctgc gccgcaggac ccccgacctg gagggcctga gggccttttt ggaggagctg 840

| | |
|--|------|
| gagttcggaa gcctcctcca cgagttcggc ctctctgggag gggagaagcc ccgggaggag | 900 |
| gccccctggc ccccgcccca aggggccttc gtgggcttcc tcctttcccg caaggagccc | 960 |
| atgtgggagg agcttctggc cctggcggcg gcctcgggag gccgcgtcca ccgggcaaca | 1020 |
| agccccgttg aggccctggc cgacctcaag gagggccggg ggttcctggc caaggacctg | 1080 |
| gccgttttgg ccctgcggga ggggggtggc ctggaccca cgacgaccc cctcctggtg | 1140 |
| gcctacctcc tggaccggc caacacccac cccgagggg tggcccgcg ctacggggg | 1200 |
| gagttcacgg aggacgcagc ggagagggcc ctctctccg agaggctctt ccagaacctc | 1260 |
| tttaaacggc tttccgagaa gctcctctgg ctctaccagg aggtggagcg gcccctctcc | 1320 |
| cgggtcttgg cccacatgga gggccgggg gtgaggctgg acgtccccct tctggaggcc | 1380 |
| ctctcctttg agctggagaa ggagatggag cgctggagg gggaggctct ccgtttggcc | 1440 |
| ggccacccct tcaacctcaa ctcccgcgac cagctggaaa gggctctctt tgacgagctg | 1500 |
| ggcctcacc cggtggggcg gacgcagaag acgggcaagc gctccaccgc ccagggggcc | 1560 |
| ctggaggccc tccggggggc ccacccatc gtggagctca tcctccagta ccgggagctt | 1620 |
| tccaagctca aaagcaccta cctggacccc ctgcccggc tcgtccacc gcggacggg | 1680 |
| cggctccaca cccgcttcaa ccagacggc acggccacgg gaaggcttcc cagctccgac | 1740 |
| cccaacctgc agaacatccc cgtgcgcacc cccttggggc agcgcatccg caaggccttc | 1800 |
| gtggccgagg aggggtggct ccttttggcg gcggactact ccagattga gctccgggtc | 1860 |
| ctggcccacc tctcggggga cgagaacctg aagcgggtct tccgggaggg gaaggacatc | 1920 |
| cataccgaga ccgcgcctg gatgttcggc ttagacccc ctctggtgga tccaaagatg | 1980 |
| cgccgggagg ccaagacggt caacttcggc gtctctacg ggatgtccg ccacaggtc | 2040 |
| tcccaggagc tcggcataga ctacaaggag gcggaggcct ttattgagcg ctacttccag | 2100 |
| agcttcccca aggtgcggg ctggatagaa aggacctgg aggaggggcg gacgcggggc | 2160 |
| tacgtggaga ccctgttcgg caggaggcg tatgtgccc acctggcctc ccgggtccgc | 2220 |
| tcggtgcggg aggcggcgga gcggatggc ttcaacatgc ccgtgcaggg caccgccc | 2280 |
| gacctgatga agatcgccat ggtcaagctc ttccccaggc taaagcccct gggggccac | 2340 |
| ctctcctcc aagtgcacaa cgagctggtc ctggagggtc ccgaggaccg ggccgaggag | 2400 |
| gccaaaggcc tggatcaagg ggtcatggag aacgcctacc ccctggacgt gcccctcgag | 2460 |
| gtggagggtg gcgtgggtcg ggactggctg gaggcgaagc aggat | 2505 |

<210> 351
 <211> 835
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 351

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Asn | Ser | Thr | Pro | Leu | Phe | Asp | Leu | Glu | Glu | Pro | Pro | Lys | Arg | Val |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Leu | Leu | Val | Asp | Gly | His | His | Leu | Ala | Tyr | Arg | Thr | Phe | Tyr | Ala | Leu |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Ser | Leu | Thr | Thr | Ser | Arg | Gly | Glu | Pro | Val | Gln | Met | Val | Tyr | Gly | Phe |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Ala | Arg | Ser | Leu | Leu | Lys | Ala | Leu | Lys | Glu | Asp | Gly | Gln | Ala | Val | Val |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| Val | Val | Phe | Asp | Ala | Lys | Ala | Pro | Ser | Phe | Arg | His | Glu | Ala | Tyr | Glu |
| 65 | | | | | 70 | | | | | 75 | | | | | 80 |
| Ala | Tyr | Lys | Ala | Gly | Arg | Ala | Pro | Thr | Pro | Glu | Asp | Phe | Pro | Arg | Gln |
| | | | | 85 | | | | | 90 | | | | | 95 | |
| Leu | Ala | Leu | Val | Lys | Arg | Leu | Val | Asp | Leu | Leu | Gly | Phe | Thr | Arg | Leu |
| | | | 100 | | | | | 105 | | | | | | 110 | |
| Glu | Ala | Pro | Gly | Tyr | Glu | Ala | Asp | Asp | Val | Leu | Gly | Thr | Leu | Ala | Lys |
| | | 115 | | | | | 120 | | | | | 125 | | | |
| Lys | Ala | Glu | Arg | Glu | Gly | Met | Glu | Val | Arg | Ile | Leu | Thr | Gly | Asp | Arg |
| | 130 | | | | | 135 | | | | | 140 | | | | |
| Asp | Phe | Phe | Gln | Leu | Leu | Ser | Glu | Lys | Val | Ser | Val | Leu | Leu | Pro | Asp |
| 145 | | | | | 150 | | | | | 155 | | | | | 160 |
| Gly | Thr | Leu | Val | Thr | Pro | Lys | Asp | Val | Gln | Glu | Lys | Tyr | Gly | Val | Pro |
| | | | | 165 | | | | | 170 | | | | | 175 | |
| Pro | Glu | Arg | Trp | Val | Asp | Phe | Arg | Ala | Leu | Thr | Gly | Asp | Arg | Ser | Asp |
| | | | 180 | | | | | 185 | | | | | | 190 | |
| Asn | Ile | Pro | Gly | Val | Ala | Gly | Ile | Gly | Glu | Lys | Thr | Ala | Leu | Arg | Leu |
| | | 195 | | | | | 200 | | | | | | 205 | | |
| Leu | Ala | Glu | Trp | Gly | Ser | Val | Glu | Asn | Leu | Leu | Lys | Asn | Leu | Asp | Arg |
| | 210 | | | | | 215 | | | | | 220 | | | | |
| Val | Lys | Pro | Asp | Ser | Leu | Arg | Arg | Lys | Ile | Glu | Ala | His | Leu | Glu | Asp |
| 225 | | | | | 230 | | | | | 235 | | | | | 240 |
| Leu | His | Leu | Ser | Leu | Asp | Leu | Ala | Arg | Ile | Arg | Thr | Asp | Leu | Pro | Leu |
| | | | | 245 | | | | | 250 | | | | | 255 | |

Glu Val Asp Phe Lys Ala Leu Arg Arg Arg Thr Pro Asp Leu Glu Gly
260 265 270

Leu Arg Ala Phe Leu Glu Glu Leu Glu Phe Gly Ser Leu Leu His Glu
275 280 285

Phe Gly Leu Leu Gly Gly Glu Lys Pro Arg Glu Glu Ala Pro Trp Pro
290 295 300

Pro Pro Glu Gly Ala Phe Val Gly Phe Leu Leu Ser Arg Lys Glu Pro
305 310 315 320

Met Trp Ala Glu Leu Leu Ala Leu Ala Ala Ala Ser Gly Gly Arg Val
325 330 335

His Arg Ala Thr Ser Pro Val Glu Ala Leu Ala Asp Leu Lys Glu Ala
340 345 350

Arg Gly Phe Leu Ala Lys Asp Leu Ala Val Leu Ala Leu Arg Glu Gly
355 360 365

Val Ala Leu Asp Pro Thr Asp Asp Pro Leu Leu Val Ala Tyr Leu Leu
370 375 380

Asp Pro Ala Asn Thr His Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly
385 390 395 400

Glu Phe Thr Glu Asp Ala Ala Glu Arg Ala Leu Leu Ser Glu Arg Leu
405 410 415

Phe Gln Asn Leu Phe Lys Arg Leu Ser Glu Lys Leu Leu Trp Leu Tyr
420 425 430

Gln Glu Val Glu Arg Pro Leu Ser Arg Val Leu Ala His Met Glu Ala
435 440 445

Arg Gly Val Arg Leu Asp Val Pro Leu Leu Glu Ala Leu Ser Phe Glu
450 455 460

Leu Glu Lys Glu Met Glu Arg Leu Glu Gly Glu Val Phe Arg Leu Ala
465 470 475 480

Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu
485 490 495

Phe Asp Glu Leu Gly Leu Thr Pro Val Gly Arg Thr Gln Lys Thr Gly
500 505 510

Lys Arg Ser Thr Ala Gln Gly Ala Leu Glu Ala Leu Arg Gly Ala His
515 520 525

Pro Ile Val Glu Leu Ile Leu Gln Tyr Arg Glu Leu Ser Lys Leu Lys
530 535 540

Ser Thr Tyr Leu Asp Pro Leu Pro Arg Leu Val His Pro Arg Thr Gly
545 550 555 560

<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 352
atgaattccc tgccctcttt tgagcccaag ggccgggtgc ttctggtgga cggccaccac 60
ctggcctacc gtaccttttt tgccctgaag ggcctcacca ccagccgcgg ggagccggtc 120
caggcgggtgt acgggttttc caagagcctt ttgaaggcgc taagggaaga cggggatgtg 180
gtgatcgtgg tgtttgacgc caaggccccc tccttcgcgc accagaccta cgaggcctac 240
aaggcggggc gggctccac ccccgaggac tttcccgcgc agcttgccct tatcaaggag 300
atggtggacc ttttgggctt taccgcctc gaggtgccgg gctttgaagc ggatgacgtc 360
ctggctaccc tggccaagaa ggcggaaaag gaaggctacg aagtgcgcat cctcaccgcg 420
gaccgggacc tttaccagct tctttcgag cgaatctcca tccttcaccc ggaggggttac 480
ctgatcacc cggagtggct ttgggagaag tatgggetta agccttccca gtgggtggac 540
taccgggcct tggccgggga cccttcgcgc aacatccccg gcgtgaaggc catcggggag 600
aagacggcgg ccaagctgat ccgggagtgg ggaagcctgg aaaaccttct taagcacctg 660
gaacaggtga aacctgcctc cgtgcgggag aagatcctta gccacatgga ggacctcaag 720
ctatccctgg agctatcccg ggtgcacacg gacttgctcc ttcagggtgga cttcgcccgg 780
cgccgggagc cggaccggga ggggcttaag gcctttttgg agaggctgga gttcgggaagc 840
ctcctccacg agttcggcct gttggaaagc ccggtggcgg cggaggaagc tccctggccg 900
cccccgagg gagccttcgt ggggtacgtt ctttcccgc cagagcccat gtgggcggag 960
cttaacgcct tggccgcgc ctggggcggc cgcgtttacc gggcgaggga tcccttgag 1020
gccttgccgg ggcttgggga ggtgaggggg cttttggcca aggacctggc ggtgctggcc 1080
ctgagggaa ggtattgcct ggcaccgggc gacgaccca tgctcctgc ctacctctg 1140
gatccttcca acaccgcccc cgaaggggta gcccggcgt acggggggga gtggaccgag 1200
gaggcggggg aaagggcgt gctttccgaa aggcctttac ccgccctcct gaagcggctt 1260
aagggggagg agaggcttct ttggctttac gaggaggtgg aaaagcccct ttcgcgggtc 1320
ctggcccaca tggaggccac gggggtacgg ttggatgtgg cctacttaaa ggccctttcc 1380
ctggaggtgg aggcggagat aaggcgcttc gaggaggagg tccaccgcct ggccgggcat 1440
cctttcaacc tgaactcccg ggaccagctg gaaaggtca tctttgacga gcttgggctt 1500

```

ccccccatcg gcaagacgca gaagacgggc aagcgctcca ccagcgccgc cgttttggag 1560
gccttgcggg aggtcatcc catcgtggac cgcaccttc agtaccggga gctttccaag 1620
ctcaaggga cctacatcga tcccttgctt gccctgggtcc accccaagac gaaccgcctc 1680
cacaccggtt tcaaccagac ggccaccgcc acggggaggc ttagcagctc ggatcctaatt 1740
ctgcaaaata tccccgtgcg cacccttttg ggccagcgga tccgcccgggc cttcgtggcc 1800
gaggaggggt ggaggctggt ggttttggac tacagccaga ttgagctcag ggtcctggcg 1860
cacctttccg gggacgagaa cctaattccg gtcttcagg agggccagga catccacacc 1920
cagacggcca gctggatggt cggcgtgcc ccagaggccg tggattccct gatgcgccgg 1980
gcggccaaga ccatcaactt cggcgtcctc tacggcatgt ccgcccaccg gctttcggga 2040
gagctggcca tcccctacga ggaggcggtg gccttcatcg agcggtatatt ccagagctac 2100
cccaagggtg gggcctggat tgagaaaacc ctggcggaag gacgggaacg gggctatgtg 2160
gaaaccctct ttggccgccc ggcgtacgtg cccgacttgg cttcccgggt gaagagcatc 2220
cgggaggcag cggagcgcat ggcttcaac atgccgttcc aggggaccgc cgcggatttg 2280
atgaaactgg ccatggtgaa gctctttccc aggcttcagg agctgggggc caggatgctt 2340
ttgcagggtg acaacgaact ggtcctcgag gctcccaagg agcaagcgga ggaagtcgcc 2400
caggaggcca agcggaccat ggaggagggtg tggccctga aggtgccctt ggagggtgaa 2460
gtgggcatcg gggaggactg gctttccgcc aaggcc 2496

```

```

<210> 353
<211> 832
<212> PRT
<213> Artificial

```

```

<220>
<223> Synthetic

```

```

<400> 353

```

```

Met Asn Ser Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu Val
1           5           10          15

Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu Lys Gly Leu
          20          25          30

Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala Lys
          35          40          45

Ser Leu Leu Lys Ala Leu Arg Glu Asp Gly Asp Val Val Ile Val Val
          50          55          60

```


| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Pro | Gly | Asp | Asp | Pro | Met | Leu | Leu | Ala | Tyr | Leu | Leu | Asp | Pro | Ser | Asn | 370 | 375 | 380 | |
| Thr | Ala | Pro | Glu | Gly | Val | Ala | Arg | Arg | Tyr | Gly | Gly | Glu | Trp | Thr | Glu | 385 | 390 | 395 | 400 |
| Glu | Ala | Gly | Glu | Arg | Ala | Leu | Leu | Ser | Glu | Arg | Leu | Tyr | Ala | Ala | Leu | 405 | 410 | 415 | |
| Leu | Lys | Arg | Leu | Lys | Gly | Glu | Glu | Arg | Leu | Leu | Trp | Leu | Tyr | Glu | Glu | 420 | 425 | 430 | |
| Val | Glu | Lys | Pro | Leu | Ser | Arg | Val | Leu | Ala | His | Met | Glu | Ala | Thr | Gly | 435 | 440 | 445 | |
| Val | Arg | Leu | Asp | Val | Ala | Tyr | Leu | Lys | Ala | Leu | Ser | Leu | Glu | Val | Glu | 450 | 455 | 460 | |
| Ala | Glu | Ile | Arg | Arg | Phe | Glu | Glu | Glu | Val | His | Arg | Leu | Ala | Gly | His | 465 | 470 | 475 | 480 |
| Pro | Phe | Asn | Leu | Asn | Ser | Arg | Asp | Gln | Leu | Glu | Arg | Val | Ile | Phe | Asp | 485 | 490 | 495 | |
| Glu | Leu | Gly | Leu | Pro | Ala | Ile | Gly | Lys | Thr | Gln | Lys | Thr | Gly | Lys | Arg | 500 | 505 | 510 | |
| Ser | Thr | Ser | Ala | Ala | Val | Leu | Glu | Ala | Leu | Arg | Glu | Ala | His | Pro | Ile | 515 | 520 | 525 | |
| Val | Asp | Arg | Ile | Leu | Gln | Tyr | Arg | Glu | Leu | Ser | Lys | Leu | Lys | Gly | Thr | 530 | 535 | 540 | |
| Tyr | Ile | Asp | Pro | Leu | Pro | Ala | Leu | Val | His | Pro | Lys | Thr | Asn | Arg | Leu | 545 | 550 | 555 | 560 |
| His | Thr | Arg | Phe | Asn | Gln | Thr | Ala | Thr | Ala | Thr | Gly | Arg | Leu | Ser | Ser | 565 | 570 | 575 | |
| Ser | Asp | Pro | Asn | Leu | Gln | Asn | Ile | Pro | Val | Arg | Thr | Pro | Leu | Gly | Gln | 580 | 585 | 590 | |
| Arg | Ile | Arg | Arg | Ala | Phe | Val | Ala | Glu | Glu | Gly | Trp | Arg | Leu | Val | Val | 595 | 600 | 605 | |
| Leu | Asp | Tyr | Ser | Gln | Ile | Glu | Leu | Arg | Val | Leu | Ala | His | Leu | Ser | Gly | 610 | 615 | 620 | |
| Asp | Glu | Asn | Leu | Ile | Arg | Val | Phe | Gln | Glu | Gly | Gln | Asp | Ile | His | Thr | 625 | 630 | 635 | 640 |
| Gln | Thr | Ala | Ser | Trp | Met | Phe | Gly | Val | Pro | Pro | Glu | Ala | Val | Asp | Ser | 645 | 650 | 655 | |
| Leu | Met | Arg | Arg | Ala | Ala | Lys | Thr | Ile | Asn | Phe | Gly | Val | Leu | Tyr | Gly | 660 | 665 | 670 | |

<213> Artificial

<220>

<223> Synthetic

<400> 356
ctcctcctcc aagtggccaa cgagctggtc ctg 33

<210> 357
<211> 33
<212> DNA
<213> Artificial

<220>

<223> Synthetic

<400> 357
caggaccagc tcgttggcca cttggaggag gag 33

<210> 358
<211> 2505
<212> DNA
<213> Artificial

<220>

<223> Synthetic

<400> 358
atgaattcca cccactttt tgacctggag gaacccccca agcgggtgct tctggtggac 60
ggccaccacc tggcctaccg caccttotat gccctgagcc tcaccacctc ccgggggggag 120
ccggtgcaga tggctctacgg ctctgccccg agcctcctca aggccttgaa ggaggacgga 180
caggcgggtgg tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacgag 240
gcctacaagg cgggccgggc ccccaacccg gaggacttcc cccgccagct cgccttggtc 300
aagcggctgg tggaccttct gggcctggtc cgcctcgagg ccccggggta cgaggcggac 360
gacgtcctgg gcaccctggc caagaaggcc gaaagggagg ggatggagggt ggcgcatcctc 420
acgggagacc gggacttctt ccagctcctc tccgagaagg tctcggtcct cctgccggac 480
gggaccctgg tcaccccaaa ggacgtccag gagaagtagc gggtgccccc ggagcgctgg 540
gtggacttcc gcgcctcac gggggacgcg tcggacaaca tccccggggt ggcggggata 600
ggggagaaga ccgcccttcg actcctcgca gagtggggga gcgtggaaaa cctcctgaag 660
aacctggacc gggtaaagcc ggactcgctc cggcgcaaga tagaggcgca cctcgaggac 720
ctccacctct ccttagacct ggccccgcatc cgcaccgacc tccccctgga ggtggacttt 780
aaggccctgc gccgcaggac ccccgacctg gagggcctga gggccttttt ggaggagctg 840

| | |
|---|------|
| gagttcggaa gcctcctcca cgagttcggc ctctctgggag gggagaagcc ccgggaggag | 900 |
| gccccctggc ccccgcccga aggggccttc gtgggcttcc tcctttcccg caaggagccc | 960 |
| atgtgggagg agcttctggc cctggcggcg gcctcgggag gccgcgtcca ccgggcaaca | 1020 |
| agccccgttg aggcctcggc cgacctcaag gagggccggg ggttcctggc caaggacctg | 1080 |
| gccgttttgg ccctgcggga gggggtggcc ctggacccca cggacgacct cctcctggtg | 1140 |
| gcctacctcc tggacctggc caacacccac cccgaggggg tggcccggcg ctacgggggc | 1200 |
| gagttcacgg aggacgcagc ggagagggcc ctctctcccg agaggctctt ccagaacctc | 1260 |
| tttaaacggc tttccgagaa gctcctctgg ctctaccagg aggtggagcg gccccctctc | 1320 |
| cgggtcttgg ccacatgga ggcccggggg gtgaggctgg acgtccccct tctggaggcc | 1380 |
| ctctcctttg agctggagaa ggagatggag cgctcggagg gggaggctct ccgtttggcc | 1440 |
| ggccacccct tcaacctcaa ctcccgcgac cagctggaaa gggctcctct tgacgagctg | 1500 |
| ggcctcacc cgggtgaagcg gacgaagaag acgggcaagc gctccaccgc ccagggggcc | 1560 |
| ctggaggccc tccggggggc ccacccatc gtggagctca tcctccagta ccgggagctt | 1620 |
| tccaagctca aaagcaccta cctggacccc ctgccccggc tcgtccacc gcggacgggc | 1680 |
| cggctccaca cccgcttcaa ccagacggcc acggccacgg gaaggctttc cagctccgac | 1740 |
| cccaacctgc agaacatccc cgtgcgcacc cccttggggc agcgcatccg caaggccttc | 1800 |
| gtggccgagg aggggtggct ccttttggcg gcggactact ccagattga gctccgggtc | 1860 |
| ctggcccacc tctcggggga cgagaacctg aagcgggtct tccgggaggg gaaggacatc | 1920 |
| cataccgaga ccgcccctg gatgttcggc ttagaccccg ctctggtgga tccaaagatg | 1980 |
| cgccgggagg ccaagacggc caacttcggc gtctctacg ggatgtccgc ccacaggctc | 2040 |
| tcccaggagc tcggcataga ctacaaggag gcggaggcct ttattgagcg ctacttccag | 2100 |
| agcttcccca aggtgcgggc ctggatagaa aggacctgg aggaggggcg gacgcggggc | 2160 |
| tacgtggaga ccctgttcgg caggaggcgc tatgtgccc acctggcctc ccgggtccgc | 2220 |
| tcggtgcggg aggcggcgga gcggatggcc ttcaacatgc ccgtgcaggg caccgcggcc | 2280 |
| gacctgatga agatcgccat ggtcaagctc ttccccaggc taaagcccct gggggcccac | 2340 |
| ctcctcctcc aagtggccaa cgagctggtc ctggagggtc ccgaggaccg ggccgaggag | 2400 |
| gccaaaggccc tgggtcaagga ggtcatggag aacgcctacc ccctggacgt gccccctgag | 2460 |
| gtggagggtg gcgtgggtcg ggactggctg gaggcgaagc aggat | 2505 |

<210> 359
 <211> 835
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 359

Met Asn Ser Thr Pro Leu Phe Asp Leu Glu Glu Pro Pro Lys Arg Val
 1 5 10 15

Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Tyr Ala Leu
 20 25 30

Ser Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Met Val Tyr Gly Phe
 35 40 45

Ala Arg Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Gln Ala Val Val
 50 55 60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu
 65 70 75 80

Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
 85 90 95

Leu Ala Leu Val Lys Arg Leu Val Asp Leu Leu Gly Leu Val Arg Leu
 100 105 110

Glu Ala Pro Gly Tyr Glu Ala Asp Asp Val Leu Gly Thr Leu Ala Lys
 115 120 125

Lys Ala Glu Arg Glu Gly Met Glu Val Arg Ile Leu Thr Gly Asp Arg
 130 135 140

Asp Phe Phe Gln Leu Leu Ser Glu Lys Val Ser Val Leu Leu Pro Asp
 145 150 155 160

Gly Thr Leu Val Thr Pro Lys Asp Val Gln Glu Lys Tyr Gly Val Pro
 165 170 175

Pro Glu Arg Trp Val Asp Phe Arg Ala Leu Thr Gly Asp Arg Ser Asp
 180 185 190

Asn Ile Pro Gly Val Ala Gly Ile Gly Glu Lys Thr Ala Leu Arg Leu
 195 200 205

Leu Ala Glu Trp Gly Ser Val Glu Asn Leu Leu Lys Asn Leu Asp Arg
 210 215 220

Val Lys Pro Asp Ser Leu Arg Arg Lys Ile Glu Ala His Leu Glu Asp
 225 230 235 240

Leu His Leu Ser Leu Asp Leu Ala Arg Ile Arg Thr Asp Leu Pro Leu

| | | | | | |
|---|-----|-----|-----|-----|-----|
| | 245 | | 250 | | 255 |
| Glu Val Asp Phe Lys Ala Leu Arg Arg Arg Thr Pro Asp Leu Glu Gly | 260 | 265 | 270 | | |
| Leu Arg Ala Phe Leu Glu Glu Leu Glu Phe Gly Ser Leu Leu His Glu | 275 | 280 | 285 | | |
| Phe Gly Leu Leu Gly Gly Glu Lys Pro Arg Glu Glu Ala Pro Trp Pro | 290 | 295 | 300 | | |
| Pro Pro Glu Gly Ala Phe Val Gly Phe Leu Leu Ser Arg Lys Glu Pro | 305 | 310 | 315 | 320 | |
| Met Trp Ala Glu Leu Leu Ala Leu Ala Ala Ala Ser Gly Gly Arg Val | 325 | 330 | 335 | | |
| His Arg Ala Thr Ser Pro Val Glu Ala Leu Ala Asp Leu Lys Glu Ala | 340 | 345 | 350 | | |
| Arg Gly Phe Leu Ala Lys Asp Leu Ala Val Leu Ala Leu Arg Glu Gly | 355 | 360 | 365 | | |
| Val Ala Leu Asp Pro Thr Asp Asp Pro Leu Leu Val Ala Tyr Leu Leu | 370 | 375 | 380 | | |
| Asp Pro Ala Asn Thr His Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly | 385 | 390 | 395 | 400 | |
| Glu Phe Thr Glu Asp Ala Ala Glu Arg Ala Leu Leu Ser Glu Arg Leu | 405 | 410 | 415 | | |
| Phe Gln Asn Leu Phe Lys Arg Leu Ser Glu Lys Leu Leu Trp Leu Tyr | 420 | 425 | 430 | | |
| Gln Glu Val Glu Arg Pro Leu Ser Arg Val Leu Ala His Met Glu Ala | 435 | 440 | 445 | | |
| Arg Gly Val Arg Leu Asp Val Pro Leu Leu Glu Ala Leu Ser Phe Glu | 450 | 455 | 460 | | |
| Leu Glu Lys Glu Met Glu Arg Leu Glu Gly Glu Val Phe Arg Leu Ala | 465 | 470 | 475 | 480 | |
| Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu | 485 | 490 | 495 | | |
| Phe Asp Glu Leu Gly Leu Thr Pro Val Lys Arg Thr Lys Lys Thr Gly | 500 | 505 | 510 | | |
| Lys Arg Ser Thr Ala Gln Gly Ala Leu Glu Ala Leu Arg Gly Ala His | 515 | 520 | 525 | | |
| Pro Ile Val Glu Leu Ile Leu Gln Tyr Arg Glu Leu Ser Lys Leu Lys | 530 | 535 | 540 | | |
| Ser Thr Tyr Leu Asp Pro Leu Pro Arg Leu Val His Pro Arg Thr Gly | | | | | |

| | | | | | | |
|------------|------------|------------|------------|------------|-------------|------|
| ctggcctacc | gtaccttttt | tgccctgaag | ggcctcacca | ccagccgcgg | ggagccggtc | 120 |
| caggcggtgt | acgggtttgc | caagagcctt | ttgaaggcgc | taagggaaga | cggggatgtg | 180 |
| gtgatcgtgg | tgtttgacgc | caaggccccc | tccttcgcc | accagaccta | cgaggcctac | 240 |
| aaggcggggc | gggctccac | ccccgaggac | ttccccggc | agcttgccct | tatcaaggag | 300 |
| atggtggacc | ttttgggcct | ggagcgcctc | gaggtgccgg | gctttgaagc | ggatgacgtc | 360 |
| ctggctaccc | tggccaagaa | ggcggaaaag | gaaggctacg | aagtgcgcat | cctcaccgcg | 420 |
| gaccgggacc | tttaccagct | tctttcggag | cgaatctcca | tccttcaccc | ggaggggttac | 480 |
| ctgatcacc | cggagtggct | ttgggagaag | tatgggctta | agccttccca | gtgggtggac | 540 |
| taccgggcct | tggccgggga | cccttcgcac | aacatccccg | gcgtgaagg | catcggggag | 600 |
| aagacggcgg | ccaagctgat | cgggagtggt | ggaagcctgg | aaaaccttct | taagcacctg | 660 |
| gaacaggtga | aacctgcctc | cgtgcgggag | aagatcctta | gccacatgga | ggacctcaag | 720 |
| ctatccctgg | agctatcccc | ggtgcacacg | gacttgctcc | ttcaggtgga | cttcgcccgg | 780 |
| cgccgggagc | cggaccggga | ggggcttaag | gcctttttgg | agaggctgga | gttcggaagc | 840 |
| ctcctccacg | agttcggcct | gttggaagc | ccggtggcgg | cggaggaagc | tccttgcccg | 900 |
| cccccgagg | gagccttcgt | ggggtacgtt | ctttcccgcc | ccgagcccat | gtgggcggag | 960 |
| cttaacgcct | tggccgcgc | ctggggcggc | cgcgtttacc | gggcggagga | tccttgagg | 1020 |
| gccttgccgg | ggcttgggga | ggtgaggggg | cttttgggca | aggacctggc | ggtgctggcc | 1080 |
| ctgagggaa | ggattgccct | ggcacccggc | gacgaccca | tgtcctcgc | ctacctctg | 1140 |
| gatccttcca | acaccgcccc | cgaaggggta | gcccggcgct | acggggggga | gtggaccgag | 1200 |
| gaggcggggg | aaagggcgct | gctttccgaa | aggctttacg | ccgccctcct | gaagcggctt | 1260 |
| aagggggagg | agaggcttct | ttggctttac | gaggaggtgg | aaaagccctt | ttcgcggtc | 1320 |
| ctggcccaca | tggaggccac | gggggtacgg | ttggatgtgg | cctacttaaa | ggccctttcc | 1380 |
| ctggaggtgg | aggcggagat | aaggcgcttc | gaggaggagg | tccacgcct | ggccgggcat | 1440 |
| cctttcaacc | tgaactcccc | ggaccagctg | gaaaggtca | tctttgacga | gcttgggctt | 1500 |
| cccgccatca | agaagacgag | gaagacgggc | aagcgctcca | ccagcgccgc | cgttttgagg | 1560 |
| gccttgccgg | aggctcatcc | catcgtggac | cgcaccttc | agtaccggga | gctttccaag | 1620 |
| ctcaagggaa | cctacatcga | tcccttgctt | gcctgggtcc | acccaagac | gaaccgcctc | 1680 |
| cacaccggtt | tcaaccagac | ggccaccgcc | acggggaggc | ttagcagctc | ggatccta | 1740 |

ctgcaaaata tccccgtgcg cacccttttg ggccagcgga tccgccgggc cttcgtggcc 1800
gaggaggggt ggaggctggt ggttttggac tacagccaga ttgagctcag ggtcctggcg 1860
cacctttccg gggacgagaa cctaataccg gtcttccagg agggccagga catccacacc 1920
cagacggcca gctggatggt cggcgtgccc ccagaggccg tggattccct gatgcgccgg 1980
gcgggccaaga ccatcaactt cggcgtcctc tacggcatgt ccgcccaccg gctttcggga 2040
gagctggcca tcccctacga ggaggcggtg gccttcacg agcggtatth ccagagctac 2100
cccaagggtgc gggcctggat tgagaaaacc ctggcggaag gacgggaacg gggctatgtg 2160
gaaaccctct ttggccgccc gcgctacgtg cccgacttgg cttcccgggt gaagagcatc 2220
cgggaggcag cggagcgcat ggccttcaac atgccgggtcc aggggaccgc cgcggatttg 2280
atgaaactgg ccatggtgaa gctctttccc aggttccagg agctgggggc caggatgctt 2340
ttgcaggtgc acaacgaact ggtcctcgag gctcccaagg agcaagcgga ggaagtcgcc 2400
caggaggcca agcggacat ggaggaggtg tggcccctga aggtgccctt ggaggtggaa 2460
gtgggcatcg gggaggactg gctttccgcc aaggcc 2496

<210> 365
<211> 832
<212> PRT
<213> Artificial

<220>
<223> Synthetic

<400> 365

Met Asn Ser Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu Val
1 5 10 15
Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu Lys Gly Leu
20 25 30
Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala Lys
35 40 45
Ser Leu Leu Lys Ala Leu Arg Glu Asp Gly Asp Val Val Ile Val Val
50 55 60
Phe Asp Ala Lys Ala Pro Ser Phe Arg His Gln Thr Tyr Glu Ala Tyr
65 70 75 80
Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu Ala
85 90 95
Leu Ile Lys Glu Met Val Asp Leu Leu Gly Leu Glu Arg Leu Glu Val
100 105 110

Pro Gly Phe Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys Lys Ala
115 120 125

Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Arg Asp Leu
130 135 140

Tyr Gln Leu Leu Ser Glu Arg Ile Ser Ile Leu His Pro Glu Gly Tyr
145 150 155 160

Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly Leu Lys Pro Ser
165 170 175

Gln Trp Val Asp Tyr Arg Ala Leu Ala Gly Asp Pro Ser Asp Asn Ile
180 185 190

Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Ala Lys Leu Ile Arg
195 200 205

Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys His Leu Glu Gln Val Lys
210 215 220

Pro Ala Ser Val Arg Glu Lys Ile Leu Ser His Met Glu Asp Leu Lys
225 230 235 240

Leu Ser Leu Glu Leu Ser Arg Val His Thr Asp Leu Leu Leu Gln Val
245 250 255

Asp Phe Ala Arg Arg Arg Glu Pro Asp Arg Glu Gly Leu Lys Ala Phe
260 265 270

Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu Leu
275 280 285

Glu Ser Pro Val Ala Ala Glu Glu Ala Pro Trp Pro Pro Pro Glu Gly
290 295 300

Ala Phe Val Gly Tyr Val Leu Ser Arg Pro Glu Pro Met Trp Ala Glu
305 310 315 320

Leu Asn Ala Leu Ala Ala Ala Trp Gly Gly Arg Val Tyr Arg Ala Glu
325 330 335

Asp Pro Leu Glu Ala Leu Arg Gly Leu Gly Glu Val Arg Gly Leu Leu
340 345 350

Ala Lys Asp Leu Ala Val Leu Ala Leu Arg Glu Gly Ile Ala Leu Ala
355 360 365

Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser Asn
370 375 380

Thr Ala Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr Glu
385 390 395 400

Glu Ala Gly Glu Arg Ala Leu Leu Ser Glu Arg Leu Tyr Ala Ala Leu
405 410 415

| | |
|---|------|
| aaggccctgc gccgcaggac ccccgacctg gagggcctga gggccttttt ggaggagctg | 840 |
| gagttcggaa gcctcctcca cgagttcggc ctctctgggag gggagaagcc ccgggaggag | 900 |
| gccccctggc ccccgccga aggggccttc gtgggcttcc tcctttcccg caaggagccc | 960 |
| atgtgggagg agcttctggc cctggcggcg gcctcgggag gccgcgtcca ccgggcaaca | 1020 |
| agccccgttg agggcctggc cgacctcaag gagggccggg ggttcctggc caaggacctg | 1080 |
| gccgttttgg ccctgcggga gggggtggcc ctggacccca cggacgaccc cctcctggtg | 1140 |
| gcctacctcc tggacccggc caacaccac cccgaggggg tggcccggcg ctacgggggc | 1200 |
| gagttcacgg aggacgcagc ggagagggcc ctctctccg agaggctctt ccagaacctc | 1260 |
| tttaaacggc tttccgagaa gctcctctgg ctctaccagg aggtggagcg gcccctctcc | 1320 |
| cgggtcttgg cccacatgga ggccccgggg gtgaggctgg acgtccccct tctggaggcc | 1380 |
| ctctcctttg agctggagaa ggagatggag cgcctggagg gggaggctct ccgtttggcc | 1440 |
| ggccacccct tcaacctcaa ctcccgcgac cagctggaaa gggtcctctt tgacgagctg | 1500 |
| ggcctcacc cggatgaagc gacgaagaag acgggcaagc gctccaccgc ccagggggcc | 1560 |
| ctggaggccc tccggggggc ccacccatc gtggagctca tcctccagta ccgggagctt | 1620 |
| tccaagctca aaagcaccta cctggacccc ctgccccggc tcgtccaccg gcggacgggc | 1680 |
| cggctccaca cccgcttcaa ccagacggcc acggccacgg gaaggctttc cagctccgac | 1740 |
| cccaacctgc agaacatccc cgtgcgcacc cccttggggc agcgcacccg caaggccttc | 1800 |
| gtggccgagg aggggtggct ccttttggcg gcggactact ccagattga gctccgggtc | 1860 |
| ctggcccacc tctcggggga cgagaacctg aagcgggtct tccgggaggg gaaggacatc | 1920 |
| cataccgaga ccgccgcctg gatgttcggc ttagaccccg ctctggtgga tccaaagatg | 1980 |
| cgcggggcgg ccaagacggt caacttcggc gtcctctacg ggatgtccgc ccacaggctc | 2040 |
| tcccaggagc tcggcataga ctacaaggag gcggaggcct ttattgagcg ctacttccag | 2100 |
| agcttcccca aggtgcgggc ctggatagaa aggaccctgg aggagggccg gacgcggggc | 2160 |
| tacgtggaga ccctgttcgg caggaggcgc tatgtgcccg acctggcctc ccgggtccgc | 2220 |
| tcggtgcggg aggcggcgga gcggatggcc ttcaacatgc ccgtgcaggg caccgccgcc | 2280 |
| gacctgatga agatcgccat ggtcaagctc ttccccaggc taaagcccct gggggccac | 2340 |
| ctctcctcc aagtggccaa cgagctggtc ctggagggtc ccgaggaccg ggccgaggag | 2400 |
| gccaaaggccc tggatcaagga ggtcatggag aacgcctacc ccctggacgt gcccctcgag | 2460 |
| gtggagggtg gcgtgggtcg ggactggctg gaggcgaagc aggat | 2505 |

<210> 367
 <211> 835
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 367

```

Met Asn Ser Thr Pro Leu Phe Asp Leu Glu Glu Pro Pro Lys Arg Val
1          5          10          15

Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Tyr Ala Leu
20        25        30

Ser Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Met Val Tyr Gly Phe
35        40        45

Ala Arg Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Gln Ala Val Val
50        55        60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu
65        70        75        80

Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
85        90        95

Leu Ala Leu Val Lys Arg Leu Val Asp Leu Leu Gly Phe Thr Arg Leu
100       105       110

Glu Ala Pro Gly Tyr Glu Ala Asp Asp Val Leu Gly Thr Leu Ala Lys
115       120       125

Lys Ala Glu Arg Glu Gly Met Glu Val Arg Ile Leu Thr Gly Asp Arg
130       135       140

Asp Phe Phe Gln Leu Leu Ser Glu Lys Val Ser Val Leu Leu Pro Asp
145       150       155       160

Gly Thr Leu Val Thr Pro Lys Asp Val Gln Glu Lys Tyr Gly Val Pro
165       170       175

Pro Glu Arg Trp Val Asp Phe Arg Ala Leu Thr Gly Asp Arg Ser Asp
180       185       190

Asn Ile Pro Gly Val Ala Gly Ile Gly Glu Lys Thr Ala Leu Arg Leu
195       200       205

Leu Ala Glu Trp Gly Ser Val Glu Asn Leu Leu Lys Asn Leu Asp Arg
210       215       220

Val Lys Pro Asp Ser Leu Arg Arg Lys Ile Glu Ala His Leu Glu Asp
225       230       235       240

```

Leu His Leu Ser Leu Asp Leu Ala Arg Ile Arg Thr Asp Leu Pro Leu
 245 250 255
 Glu Val Asp Phe Lys Ala Leu Arg Arg Arg Thr Pro Asp Leu Glu Gly
 260 265 270
 Leu Arg Ala Phe Leu Glu Glu Leu Glu Phe Gly Ser Leu Leu His Glu
 275 280 285
 Phe Gly Leu Leu Gly Gly Glu Lys Pro Arg Glu Glu Ala Pro Trp Pro
 290 295 300
 Pro Pro Glu Gly Ala Phe Val Gly Phe Leu Leu Ser Arg Lys Glu Pro
 305 310 315 320
 Met Trp Ala Glu Leu Leu Ala Leu Ala Ala Ala Ser Gly Gly Arg Val
 325 330 335
 His Arg Ala Thr Ser Pro Val Glu Ala Leu Ala Asp Leu Lys Glu Ala
 340 345 350
 Arg Gly Phe Leu Ala Lys Asp Leu Ala Val Leu Ala Leu Arg Glu Gly
 355 360 365
 Val Ala Leu Asp Pro Thr Asp Asp Pro Leu Leu Val Ala Tyr Leu Leu
 370 375 380
 Asp Pro Ala Asn Thr His Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly
 385 390 395 400
 Glu Phe Thr Glu Asp Ala Ala Glu Arg Ala Leu Leu Ser Glu Arg Leu
 405 410 415
 Phe Gln Asn Leu Phe Lys Arg Leu Ser Glu Lys Leu Leu Trp Leu Tyr
 420 425 430
 Gln Glu Val Glu Arg Pro Leu Ser Arg Val Leu Ala His Met Glu Ala
 435 440 445
 Arg Gly Val Arg Leu Asp Val Pro Leu Leu Glu Ala Leu Ser Phe Glu
 450 455 460
 Leu Glu Lys Glu Met Glu Arg Leu Glu Gly Glu Val Phe Arg Leu Ala
 465 470 475 480
 Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu
 485 490 495
 Phe Asp Glu Leu Gly Leu Thr Pro Val Lys Arg Thr Lys Lys Thr Gly
 500 505 510
 Lys Arg Ser Thr Ala Gln Gly Ala Leu Glu Ala Leu Arg Gly Ala His
 515 520 525
 Pro Ile Val Glu Leu Ile Leu Gln Tyr Arg Glu Leu Ser Lys Leu Lys
 530 535 540

<210> 368
 <211> 2496
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 368
 atgaattccc tgccctctt tgagcccaag ggccgggtgc ttctggtgga cggccaccac 60
 ctggcctacc gtacctttt tgccctgaag ggcctcacca ccagccgcgg ggagccggtc 120
 caggcgggtg acgggtttgc caagagcctt ttgaaggcgc taagggaaga cggggatgtg 180
 gtgatcgtgg tgtttgacgc caaggccccc tccttcgcgc accagaccta cgaggcctac 240
 aaggcggggc gggctccac ccccgaggac tttcccggc agcttgccct tatcaaggag 300
 atggtggacc ttttgggctt taccgcctc gaggtgccgg gctttgaagc ggatgacgtc 360
 ctggctaccc tggccaagaa ggcggaaaag gaaggctacg aagtgcgcac cctcaccgcg 420
 gaccgggacc tttaccagct tctttcggag cgaatctcca tccttcaccc ggaggggttac 480
 ctgatcacc cggagtggct ttgggagaag tatgggctta agccttccca gtgggtggac 540
 taccgggcct tggccgggga cccttcgcac aacatccccg gcgtgaaggg catcggggag 600
 aagacggcgg ccaagctgat ccgggagtgg ggaagcctgg aaaaccttct taagcacctg 660
 gaacaggtga aacctgcctc cgtgcgggag aagatcctta gccacatgga ggacctcaag 720
 ctatccctgg agctatcccg ggtgcacacg gacttgetcc ttcaggtgga cttcgcccg 780
 cgccgggagc cggaccggga ggggcttaag gcctttttgg agaggctgga gttcgggaagc 840
 ctccctcacg agttcggcct gttggaaagc ccggtggcgg cggaggaagc tccctggccg 900
 cccccgagg gagccttcgt ggggtacgtt ctttcccgcc ccgagcccat gtgggaggag 960
 cttaacgcct tggccgccc ctggggcggc cgcgtttacc gggcggagga tcccttgag 1020
 gccttgccgg ggcttgggga ggtgaggggg cttttggcca aggacctggc ggtgctggcc 1080
 ctgagggaa ggattgccct ggcaccgggc gacgaccca tgctcctcgc ctacctctg 1140
 gatccttcca acaccgcccc cgaaggggta gcccggcgt acggggggga gtggaccgag 1200
 gaggcggggg aaaggcgct gctttccgaa aggctttacg ccgccctcct gaagcggcct 1260
 aagggggagg agaggcttct ttggctttac gaggaggtgg aaaagccctt ttcgcgggtc 1320
 ctggcccaca tggaggccac gggggtacgg ttggatgtgg cctacttaaa ggccctttcc 1380
 ctggaggtgg aggcggagat aaggcgcttc gaggaggagg tccaccgcct ggccgggcat 1440

| | | |
|---|-----|-----|
| 660 | 665 | 670 |
| Met Ser Ala His Arg Leu Ser Gly Glu Leu Ala Ile Pro Tyr Glu Glu | | |
| 675 | 680 | 685 |
| Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Tyr Pro Lys Val Arg | | |
| 690 | 695 | 700 |
| Ala Trp Ile Glu Lys Thr Leu Ala Glu Gly Arg Glu Arg Gly Tyr Val | | |
| 705 | 710 | 715 |
| Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Ala Ser Arg | | |
| 725 | 730 | 735 |
| Val Lys Ser Ile Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro | | |
| 740 | 745 | 750 |
| Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu | | |
| 755 | 760 | 765 |
| Phe Pro Arg Leu Gln Glu Leu Gly Ala Arg Met Leu Leu Gln Val His | | |
| 770 | 775 | 780 |
| Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Gln Ala Glu Glu Val Ala | | |
| 785 | 790 | 795 |
| Gln Glu Ala Lys Arg Thr Met Glu Glu Val Trp Pro Leu Lys Val Pro | | |
| 805 | 810 | 815 |
| Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys Ala | | |
| 820 | 825 | 830 |

<210> 370
 <211> 33
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 370
 ccctccgaca acctcgccgg ggtcaagggc atc

33

<210> 371
 <211> 33
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 371
 ccctccgaca acctcaaggg ggtcaagggc atc

33

<210> 372

<211> 18
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 372
gaggttgctc gaggggtc

18

<210> 373
<211> 2526
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 373
atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
gaaccggtgc aggcggtcta cggtctcgcc aagagcctcc tcaaggccct gaaggaggac 180
gggtacaagg ccgtcttcgt ggtctttgac gccaaagccc cctccttcg ccacgaggcc 240
tacgaggcct acaaggcggg gagggccccg acccccgagg acttccccg gcagctcgcc 300
ctcatcaagg agctggtgga cctcctgggg tttaccgcc tcgaggctcc cggtacgag 360
gaggacgacg ttctcgccac cctggccaag aaggcggaag aggaggggta cgaggtgcgc 420
atcctcaccg ccgaccgga cctctacca ctcgtctccg accgcgtcgc cgtcctccac 480
cccagaggcc acctcatcac ccgggagtgg ctttgggaga agtacggcct caggccggag 540
cagtgggtgg acttccgcgc cctcgtgggg gaccctccg acaacctcgc cgggggtcaag 600
ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc 660
ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg 720
gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctccc cctggagggtg 780
gacctcgccc aggggcggga gcccgaaccg gaggggctta gggccttcct ggagaggctg 840
gagttcgga gcctcctcca cgagttcggc ctcttgagg ccccgcccc cctggaggag 900
gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctcccg ccccgagccc 960
atgtgggcgg agcttaaagc cctggccgcc tgcaggggag gccgcgtgca ccgggcagca 1020
gacccttgag cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc 1080
gccgtcttgg cctcgaggga ggggctagac ctcgtgccc gggacgaccc catgctcctc 1140

```

gcctacctcc tggacccttc gaacaccacc cccgaggggg tggcgcggcg ctacgggggg 1200
gagtggacgg aggacgccgc ccaccgggcc ctctctctcg agaggctcca tcggaacctc 1260
cttaagcgcc tcgaggggga ggagaagctc ctttggctct accacgaggt ggaaaagccc 1320
ctctcccggg tcctggccca tatggaggcc accggggtac ggcgggacgt ggcctacctt 1380
caggcccttt ccctggagct tgcggaggag atccgcccgc tcgaggagga ggtcttccgc 1440
ttggcgggcc accccttcaa cctcaactcc cgggaccagc tggaaagggg gctctttgac 1500
gagcttaggc ttccgcctt gaagaagacg aagaagacag gcaagcgctc caccagcgcc 1560
gcggtgctgg aggccctacg ggaggccac cccatcgtag agaagatcct ccagcaccgg 1620
gagctcacca agctcaagaa cacctacgtg gacccctcc caagcctcgt ccaccgagg 1680
acgggcccgc tccacaccg cttcaaccag acggccacgg ccacggggag gcttagtagc 1740
tccgaccca acctgcagaa catccccgtc cgcacccct tgggccagag gatccgccc 1800
gccttcgtgg ccgaggcggg ttgggcttg gtggccctgg actatagcca gatagagctc 1860
cgcttcctcg cccacctctc cggggacgaa aacctgatca ggtcttcca ggaggggaag 1920
gacatccaca cccagaccgc aagctggatg ttcggcgctc ccccgaggc cgtggacccc 1980
ctgatgcgcc gggcggccaa gacggtgaac ttcggcgctc tctacggcat gtccgccc 2040
aggctctccc aggagcttgc catccccctac gaggaggcgg tggcctttat agagcgctac 2100
ttccaaagct tccccaaagt gcgggcctgg atagaaaaga ccctggagga ggggaggaag 2160
cggggctacg tggaaaccct cttcggaaga aggcgctacg tgcccgcact caacgcccgg 2220
gtgaagagcg tcaggaggc cgcgagcgc atggccttca acatgcccgt ccagggcacc 2280
gccgcccacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg 2340
gcccgcacgc tcctccaggt cgccaacgag ctctctctgg agggcccca agcgcgggcc 2400
gaggaggtgg cggttttggc caaggaggcc atggagaagg cctatcccct cgccgtgccc 2460
ctggaggtgg aggtggggat gggggaggac tggctttccg ccaaggggtca ccaccaccac 2520
caccac 2526

```

<210> 374
 <211> 842
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 374

```

Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
1          5          10          15

Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu
20          25          30

Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
35          40          45

Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala
50          55          60

Val Phe Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala
65          70          75          80

Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro
85          90          95

Arg Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr
100         105         110

Arg Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu
115         120         125

Ala Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala
130         135         140

Asp Arg Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His
145         150         155         160

Pro Glu Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly
165         170         175

Leu Arg Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro
180         185         190

Ser Asp Asn Leu Ala Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu
195         200         205

Lys Leu Leu Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu
210         215         220

Asp Arg Val Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu
225         230         235         240

Glu Asp Leu Arg Leu Ser Leu Glu Leu Ser Arg Val Arg Thr Asp Leu
245         250         255

Pro Leu Glu Val Asp Leu Ala Gln Gly Arg Glu Pro Asp Arg Glu Gly
260         265         270

Leu Arg Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu
275         280         285

```


Phe Gly Leu Leu Glu Ala Pro Ala Pro Leu Glu Glu Ala Pro Trp Pro
290 295 300

Pro Pro Glu Gly Ala Phe Val Gly Phe Val Leu Ser Arg Pro Glu Pro
305 310 315 320

Met Trp Ala Glu Leu Lys Ala Leu Ala Ala Cys Arg Gly Gly Arg Val
325 330 335

His Arg Ala Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val
340 345 350

Arg Gly Leu Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly
355 360 365

Leu Asp Leu Val Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu
370 375 380

Asp Pro Ser Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly
385 390 395 400

Glu Trp Thr Glu Asp Ala Ala His Arg Ala Leu Leu Ser Glu Arg Leu
405 410 415

His Arg Asn Leu Leu Lys Arg Leu Glu Gly Glu Glu Lys Leu Leu Trp
420 425 430

Leu Tyr His Glu Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met
435 440 445

Glu Ala Thr Gly Val Arg Arg Asp Val Ala Tyr Leu Gln Ala Leu Ser
450 455 460

Leu Glu Leu Ala Glu Glu Ile Arg Arg Leu Glu Glu Glu Val Phe Arg
465 470 475 480

Leu Ala Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg
485 490 495

Val Leu Phe Asp Glu Leu Arg Leu Pro Ala Leu Lys Lys Thr Lys Lys
500 505 510

Thr Gly Lys Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu
515 520 525

Ala His Pro Ile Val Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys
530 535 540

Leu Lys Asn Thr Tyr Val Asp Pro Leu Pro Ser Leu Val His Pro Arg
545 550 555 560

Thr Gly Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly
565 570 575

Arg Leu Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr
580 585 590


```

tccgacccca acctgcagaa catccccgtc cgcaccccct tgggccagag gatccgccgg 1800
gccttcgtgg ccgaggcggg ttgggcgttg gtggcccttg actatagcca gatagagctc 1860
cgcgctctcg cccacctctc cggggacgaa aacctgatca gggctctcca ggaggggaag 1920
gacatccaca cccagaccgc aagctggatg ttcggcgctc ccccgagggc cgtggacccc 1980
ctgatgcgcc gggcggccaa gacggtgaac ttcggcgctc tctacggcat gtccgcccac 2040
aggctctccc aggagcttgc catcccctac gaggaggcgg tggcctttat agagcgctac 2100
ttccaaagct tccccaaagt gcgggccttg atagaaaaga ccctggagga ggggaggaag 2160
cggggctacg tggaaacctt cttcggaaga aggcgctacg tgcccgacct caacgcccgg 2220
gtgaagagcg tcaggagggc cgcggagcgc atggccttca acatgcccgt ccagggcacc 2280
gccgccgacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg 2340
gcccgcacgc tcctccaggt cgccaacgag ctctctcttg agggccccca agcgcgggcc 2400
gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatcccct cgccgtgccc 2460
ctggaggtgg aggtggggat gggggaggac tggctttccg ccaagggtca ccaccaccac 2520
caccac 2526

```

```

<210> 376
<211> 842
<212> PRT
<213> Artificial

```

```

<220>
<223> Synthetic

```

```

<400> 376

```

```

Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
1          5          10          15
Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu
20          25          30
Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
35          40          45
Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala
50          55          60
Val Phe Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala
65          70          75          80
Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro
85          90          95

```


| | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Glu | Trp | Thr | Glu | Asp | Ala | Ala | His | Arg | Ala | Leu | Leu | Ser | Glu | Arg | Leu | 405 | 410 | 415 |
| His | Arg | Asn | Leu | Leu | Lys | Arg | Leu | Glu | Gly | Glu | Glu | Lys | Leu | Leu | Trp | 420 | 425 | 430 |
| Leu | Tyr | His | Glu | Val | Glu | Lys | Pro | Leu | Ser | Arg | Val | Leu | Ala | His | Met | 435 | 440 | 445 |
| Glu | Ala | Thr | Gly | Val | Arg | Arg | Asp | Val | Ala | Tyr | Leu | Gln | Ala | Leu | Ser | 450 | 455 | 460 |
| Leu | Glu | Leu | Ala | Glu | Glu | Ile | Arg | Arg | Leu | Glu | Glu | Glu | Val | Phe | Arg | 465 | 470 | 475 |
| Leu | Ala | Gly | His | Pro | Phe | Asn | Leu | Asn | Ser | Arg | Asp | Gln | Leu | Glu | Arg | 485 | 490 | 495 |
| Val | Leu | Phe | Asp | Glu | Leu | Arg | Leu | Pro | Ala | Leu | Lys | Lys | Thr | Lys | Lys | 500 | 505 | 510 |
| Thr | Gly | Lys | Arg | Ser | Thr | Ser | Ala | Ala | Val | Leu | Glu | Ala | Leu | Arg | Glu | 515 | 520 | 525 |
| Ala | His | Pro | Ile | Val | Glu | Lys | Ile | Leu | Gln | His | Arg | Glu | Leu | Thr | Lys | 530 | 535 | 540 |
| Leu | Lys | Asn | Thr | Tyr | Val | Asp | Pro | Leu | Pro | Ser | Leu | Val | His | Pro | Arg | 545 | 550 | 555 |
| Thr | Gly | Arg | Leu | His | Thr | Arg | Phe | Asn | Gln | Thr | Ala | Thr | Ala | Thr | Gly | 565 | 570 | 575 |
| Arg | Leu | Ser | Ser | Ser | Asp | Pro | Asn | Leu | Gln | Asn | Ile | Pro | Val | Arg | Thr | 580 | 585 | 590 |
| Pro | Leu | Gly | Gln | Arg | Ile | Arg | Arg | Ala | Phe | Val | Ala | Glu | Ala | Gly | Trp | 595 | 600 | 605 |
| Ala | Leu | Val | Ala | Leu | Asp | Tyr | Ser | Gln | Ile | Glu | Leu | Arg | Val | Leu | Ala | 610 | 615 | 620 |
| His | Leu | Ser | Gly | Asp | Glu | Asn | Leu | Ile | Arg | Val | Phe | Gln | Glu | Gly | Lys | 625 | 630 | 635 |
| Asp | Ile | His | Thr | Gln | Thr | Ala | Ser | Trp | Met | Phe | Gly | Val | Pro | Pro | Glu | 645 | 650 | 655 |
| Ala | Val | Asp | Pro | Leu | Met | Arg | Arg | Ala | Ala | Lys | Thr | Val | Asn | Phe | Gly | 660 | 665 | 670 |
| Val | Leu | Tyr | Gly | Met | Ser | Ala | His | Arg | Leu | Ser | Gln | Glu | Leu | Ala | Ile | 675 | 680 | 685 |
| Pro | Tyr | Glu | Glu | Ala | Val | Ala | Phe | Ile | Glu | Arg | Tyr | Phe | Gln | Ser | Phe | 690 | 695 | 700 |

Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys
705 710 715 720

Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp
725 730 735

Leu Asn Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala
740 745 750

Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala
755 760 765

Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu
770 775 780

Leu Gln Val Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala
785 790 795 800

Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro
805 810 815

Leu Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu
820 825 830

Ser Ala Lys Gly His His His His His His
835 840

<210> 377
<211> 36
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 377
gagaggctcc atcggaagaa gcttaagcgc ctcgag

36

<210> 378
<211> 18
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 378
ccgatggagc ctctccga

18

<210> 379
<211> 2526
<212> DNA
<213> Artificial'

<220>

<223> Synthetic

<400> 379

atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
gaaccggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac 180
gggtacaagg ccgtcttcgt ggtctttgac gccaaaggccc cctccttccg ccacgaggcc 240
tacgaggcct acaaggcggg gagggccccc acccccaggg acttcccccg gcagctcgcc 300
ctcatcaagg agctggtgga cctcctgggg ttaccgccg tcgagggtccc cggctacgag 360
gcggacgacg ttctcgccac cctggccaag aaggcgaaa aggaggggta cgagggtgcgc 420
atcctcaccg ccgaccgca cctctaccaa ctcgctctcc accgcgtcgc cgtcctccac 480
cccagaggcc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag 540
cagtgggtgg acttccgcgc cctcgtgggg gacctctccg acaacctccc cggggtcaag 600
ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc 660
ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg 720
gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctccc cctggagggtg 780
gacctcggcc aggggcggga gcccgaaccg gaggggctta gggccttcct ggagaggctg 840
gagttcggca gcctcctcca cgagttcggc ctcttgagg ccccgcccc cctggaggag 900
gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctccc ccccgagccc 960
atgtgggcgg agcttaaagc cctggccgcc tgcaggggcg gccgcgtgca ccgggcagca 1020
gacctcttgg cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc 1080
gccgtcttgg cctcgaggga ggggctagac ctcgtgccc gggacgacct catgctcctc 1140
gcctacctcc tggaccttc gaacaccacc cccgaggggg tggcgcggg ctacgggggg 1200
gagtggacgg aggacgccgc ccaccgggcc ctctctcgg agaggctcca tcggaagaag 1260
cttaagcgcc tcgaggggga ggagaagctc ctttggctct accacgaggt ggaaaagccc 1320
ctctcccggg tcctggccca tatggaggcc accggggtac ggcgggacgt ggcctacctt 1380
caggcccttt ccctggagct tgcggaggag atccgccgcc tcgaggagga ggtcttccgc 1440
ttggcggggc accccttcaa cctcaactcc cgggaccagc tggaaagggt gctctttgac 1500
gagcttaggc ttcccgctt gaagaagacg aagaagacag gcaagcgctc caccagcgcc 1560
gcggtgctgg aggcctacg ggaggccac cccatcgtgg agaagatcct ccagcaccgg 1620

gagctcacca agctcaagaa cacctacgtg gacccccctcc caagcctcgt ccacccgagg 1680
acggggccgcc tccacacccg cttcaaccag acggccacgg ccacggggag gcttagtagc 1740
tccgacccca acctgcagaa catccccgtc cgcacccccct tgggccagag gatccgccgg 1800
gccttcgtgg ccgaggcggg ttgggcgttg gtggcccttg actatagcca gatagagctc 1860
cgcgtcctcg cccacctctc cggggacgaa aacctgatca gggctctcca ggaggggaag 1920
gacatccaca cccagaccgc aagctggatg ttcggcgtcc ccccggaggc cgtggacccc 1980
ctgatgcgcc gggcggccaa gacggtgaac ttcggcgtcc tctacggcat gtccgcccat 2040
aggtctctcc aggagcttgc catcccctac gaggaggcgg tggcctttat agagcgctac 2100
ttccaaagct tccccaaagt gcgggccttg atagaaaaga ccttggagga ggggaggaag 2160
cggggctacg tggaaaccct cttcggaaga aggcgctacg tgcccgaact caacgcccgg 2220
gtgaagagcg tcaggagggc cgcggagcgc atggccttca acatgcccggt ccagggcacc 2280
gccgccgacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg 2340
gcccgcacgc tctccaggt cgccaacgag ctctccttg aggcccccca agcgcggggc 2400
gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatcccct cgcctgccc 2460
ctggaggtgg aggtggggat gggggaggac tggctttccg ccaagggtca ccaccaccac 2520
caccac 2526

<210> 380
<211> 842
<212> PRT
<213> Artificial

<220>
<223> Synthetic

<400> 380

Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
1 5 10 15

Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu
20 25 30

Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
35 40 45

Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala
50 55 60

Val Phe Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala

| | | |
|---|-----|-------------|
| 370 | 375 | 380 |
| Asp Pro Ser Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly | | |
| 385 | 390 | 395 400 |
| Glu Trp Thr Glu Asp Ala Ala His Arg Ala Leu Leu Ser Glu Arg Leu | | |
| | 405 | 410 415 |
| His Arg Lys Lys Leu Lys Arg Leu Glu Gly Glu Glu Lys Leu Leu Trp | | |
| | 420 | 425 430 |
| Leu Tyr His Glu Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met | | |
| | 435 | 440 445 |
| Glu Ala Thr Gly Val Arg Arg Asp Val Ala Tyr Leu Gln Ala Leu Ser | | |
| | 450 | 455 460 |
| Leu Glu Leu Ala Glu Glu Ile Arg Arg Leu Glu Glu Glu Val Phe Arg | | |
| | 465 | 470 475 480 |
| Leu Ala Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg | | |
| | 485 | 490 495 |
| Val Leu Phe Asp Glu Leu Arg Leu Pro Ala Leu Lys Lys Thr Lys Lys | | |
| | 500 | 505 510 |
| Thr Gly Lys Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu | | |
| | 515 | 520 525 |
| Ala His Pro Ile Val Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys | | |
| | 530 | 535 540 |
| Leu Lys Asn Thr Tyr Val Asp Pro Leu Pro Ser Leu Val His Pro Arg | | |
| | 545 | 550 555 560 |
| Thr Gly Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly | | |
| | 565 | 570 575 |
| Arg Leu Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr | | |
| | 580 | 585 590 |
| Pro Leu Gly Gln Arg Ile Arg Arg Ala Phe Val Ala Glu Ala Gly Trp | | |
| | 595 | 600 605 |
| Ala Leu Val Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala | | |
| | 610 | 615 620 |
| His Leu Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys | | |
| | 625 | 630 635 640 |
| Asp Ile His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu | | |
| | 645 | 650 655 |
| Ala Val Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly | | |
| | 660 | 665 670 |
| Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile | | |

| 675 | | | | | 680 | | | | | 685 | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Pro | Tyr | Glu | Glu | Ala | Val | Ala | Phe | Ile | Glu | Arg | Tyr | Phe | Gln | Ser | Phe |
| 690 | | | | | | 695 | | | | | 700 | | | | |
| Pro | Lys | Val | Arg | Ala | Trp | Ile | Glu | Lys | Thr | Leu | Glu | Glu | Gly | Arg | Lys |
| 705 | | | | | | 710 | | | | | 715 | | | | 720 |
| Arg | Gly | Tyr | Val | Glu | Thr | Leu | Phe | Gly | Arg | Arg | Arg | Tyr | Val | Pro | Asp |
| | | | | 725 | | | | | 730 | | | | | 735 | |
| Leu | Asn | Ala | Arg | Val | Lys | Ser | Val | Arg | Glu | Ala | Ala | Glu | Arg | Met | Ala |
| | | | 740 | | | | | 745 | | | | | 750 | | |
| Phe | Asn | Met | Pro | Val | Gln | Gly | Thr | Ala | Ala | Asp | Leu | Met | Lys | Leu | Ala |
| | | 755 | | | | | 760 | | | | | 765 | | | |
| Met | Val | Lys | Leu | Phe | Pro | Arg | Leu | Arg | Glu | Met | Gly | Ala | Arg | Met | Leu |
| | | 770 | | | | 775 | | | | | 780 | | | | |
| Leu | Gln | Val | Ala | Asn | Glu | Leu | Leu | Leu | Glu | Ala | Pro | Gln | Ala | Arg | Ala |
| 785 | | | | | | 790 | | | | | 795 | | | | 800 |
| Glu | Glu | Val | Ala | Ala | Leu | Ala | Lys | Glu | Ala | Met | Glu | Lys | Ala | Tyr | Pro |
| | | | 805 | | | | | 810 | | | | | | 815 | |
| Leu | Ala | Val | Pro | Leu | Glu | Val | Glu | Val | Gly | Met | Gly | Glu | Asp | Trp | Leu |
| | | | 820 | | | | 825 | | | | | | 830 | | |
| Ser | Ala | Lys | Gly | His | His | His | His | His | His | | | | | | |
| | | 835 | | | | | 840 | | | | | | | | |

<210> 381
 <211> 33
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 381
 gtgcgcaccg acctcctcct ggaggtggac ctc

33

<210> 382
 <211> 33
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 382
 gaggtccacc tccaggagga ggtcgggtgcg cac

33

<210> 383

<211> 2526
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 383
 atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
 ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
 gaaccggtgc aggcggtcta cggttcgcc aagagcctcc tcaaggccct gaaggaggac 180
 ggggtacaagg ccgtcttcgt ggtctttgac gccaaaggccc cctccttcg ccacgaggcc 240
 tacgaggcct acaaggcggg gagggccccg acccccgagg acttcccccg gcagctcgcc 300
 ctcatcaagg agctggtgga cctcctgggg ttaccgccg tcgagggtccc cggctacgag 360
 gcggacgacg ttctcgccac cctggccaag aaggcggaag aggaggggta cgagggtgcgc 420
 atcctcaccg ccgaccgcga cctctaccaa ctcgctctcc accgcgtcgc cgtcctccac 480
 cccgagggcc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag 540
 cagtgggtgg acttccgcgc cctcgtgggg gaccctccg acaacctccc cgggggtcaag 600
 ggcacgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc 660
 ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg 720
 gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctcct cctggagggtg 780
 gacctcgccc aggggcgggg gcccgaccgg gaggggctta gggccttcct ggagaggctg 840
 gagttcggca gcctcctcca cgagttcggc ctcttgagg ccccgcccc cctggaggag 900
 gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctcccg ccccgagccc 960
 atgtgggcgg agcttaaagc cctggccgcc tgcaggggcg gccgcgtgca ccgggcagca 1020
 gacccttgg cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc 1080
 gccgtcttgg cctcgaggga ggggctagac ctcgtgcccg gggacgacct catgctcctc 1140
 gcctacctcc tggacccttc gaacaccacc cccgaggggg tggcgcgggc ctacgggggg 1200
 gagtggacgg aggacgcgc ccaccgggcc ctctctcgg agaggctcca tcggaacctc 1260
 cttaagcgcc tcgaggggga ggagaagctc ctttggtctt accacgaggt ggaaaagccc 1320
 ctctcccggg tcctggccca tatggaggcc accggggtac ggcgggacgt ggcctacctt 1380
 caggcccttt ccctggagct tgcggaggag atccgccgcc tcgaggagga ggtcttccgc 1440
 ttggcgggcc accccttcaa cctcaactcc cgggaccagc tggaaagggt gctctttgac 1500

09364636-052404

Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala
 50 55 60
 Val Phe Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala
 65 70 75 80
 Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro
 85 90 95
 Arg Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr
 100 105 110
 Arg Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu
 115 120 125
 Ala Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala
 130 135 140
 Asp Arg Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His
 145 150 155 160
 Pro Glu Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly
 165 170 175
 Leu Arg Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro
 180 185 190
 Ser Asp Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu
 195 200 205
 Lys Leu Leu Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu
 210 215 220
 Asp Arg Val Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu
 225 230 235 240
 Glu Asp Leu Arg Leu Ser Leu Glu Leu Ser Arg Val Arg Thr Asp Leu
 245 250 255
 Leu Leu Glu Val Asp Leu Ala Gln Gly Arg Glu Pro Asp Arg Glu Gly
 260 265 270
 Leu Arg Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu
 275 280 285
 Phe Gly Leu Leu Glu Ala Pro Ala Pro Leu Glu Glu Ala Pro Trp Pro
 290 295 300
 Pro Pro Glu Gly Ala Phe Val Gly Phe Val Leu Ser Arg Pro Glu Pro
 305 310 315 320
 Met Trp Ala Glu Leu Lys Ala Leu Ala Ala Cys Arg Gly Gly Arg Val
 325 330 335
 His Arg Ala Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val
 340 345 350

Ala Val Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly
660 665 670

Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile
675 680 685

Pro Tyr Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe
690 695 700

Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys
705 710 715 720

Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp
725 730 735

Leu Asn Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala
740 745 750

Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala
755 760 765

Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu
770 775 780

Leu Gln Val Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala
785 790 795 800

Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro
805 810 815

Leu Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu
820 825 830

Ser Ala Lys Gly His His His His His His
835 840

<210> 385
<211> 33
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 385
ggggccttcg tgggctacgt cctctccgc ccc

33

<210> 386
<211> 33
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 386

ggggcgggag aggacgtagc ccacgaaggc ccc

33

<210> 387
<211> 2526
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 387
atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
gaaccggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac 180
gggtacaagg ccgtcttcgt ggtctttgac gccaaaggccc cctccttcgg ccacgaggcc 240
tacgaggcct acaaggcggg gagggccccg acccccgagg acttcccccg gcagctcgcc 300
ctcatcaagg agctggtgga cctcctgggg tttaccgcc tccaggtccc cggctacgag 360
gcggaacgacg ttctcgccac cctggccaag aaggcggaag aggaggggta cgaggtgcgc 420
atcctcaccg ccgaccgca cctctaccaa ctgctctccg accgctgcgc cgtcctccac 480
cccaggggcc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag 540
cagtgggtgg acttccgcgc cctcgtgggg gaccctccg acaacctccc cggggtcaag 600
ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc 660
ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg 720
gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctccc cctggagggtg 780
gacctcgccc aggggcggga gcccgaccgg gaggggctta gggccttcct ggagaggctg 840
gagttcggca gcctcctcca cgagttcggc ctcttgaggg ccccgcccc cctggaggag 900
gccccctggc ccccgccgga aggggccttc gtgggctacg tcctctcccg ccccgagccc 960
atgtgggcgg agcttaaagc cctggccgcc tgcaggggag gccgctgca ccgggcagca 1020
gacccttgg cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc 1080
gccgtcttgg cctcgaggga ggggctagac ctctgcccc gggacgaccc catgctcctc 1140
gcctacctcc tggaccttc gaacaccacc cccgaggggg tggcgcggcg ctacgggggg 1200
gagtggacgg aggacgccgc ccaccgggcc ctctctcgg agaggctcca tcggaacctc 1260
cttaagcgcc tcgaggggga ggagaagctc ctttggctct accacgaggt ggaaaagccc 1320
ctctccccgg tcctggccca tatggaggcc accggggtac ggcgggacgt ggcctacctt 1380

caggcccttt ccctggagct tgcggaggag atccgccgcc tcgaggagga ggtcttccgc 1440
 ttggcggggcc accccttcaa cctcaactcc cgggaccagc tggaaagggg gctctttgac 1500
 gagcttaggc ttcccgctt gaagaagacg aagaagacag gcaagcgctc caccagcgcc 1560
 gcggtgctgg aggccttacg ggaggccac cccatcgtag agaagatcct ccagcaccgg 1620
 gagctcacca agtcaagaa cacctacgtg gacccctcc caagcctcgt ccaccgagg 1680
 acgggccgcc tccacaccg cttcaaccag acggccacgg ccacggggag gcttagtagc 1740
 tccgaccca acctgcagaa catccccgtc cgcacccct tgggccagag gatccgccgg 1800
 gccttcgtgg ccgaggcggg ttgggcgttg gtggccctgg actatagcca gatagagctc 1860
 cgcgtcctcg ccacctctc cggggacgaa aacctgatca ggtcttcca ggaggggaag 1920
 gacatccaca ccagaccgc aagctggatg ttcggcgctc ccccgagggc cgtggacccc 1980
 ctgatgcgcc gggcgccaa gacggtgaac ttcggcgctc tctacggcat gtccgcccat 2040
 aggtctctcc aggagcttgc catcccctac gaggaggcgg tggcctttat agagcgctac 2100
 ttccaaagct tccccaaagt gcgggcctgg atagaaaaga ccctggagga ggggaggaag 2160
 cggggctacg tggaaaccct cttcggaaga aggcgctac tgcccgacct caacgcccgg 2220
 gtgaagagcg tcaggagggc cgcggagcgc atggccttca acatgcccgt ccagggcacc 2280
 gccgccgacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg 2340
 gcccgcatgc tcctccaggt cgccaacgag ctctctctgg agggccccc aagcggggcc 2400
 gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatcccct cgccgtgccc 2460
 ctggaggtgg aggtggggat gggggaggac tggctttccg ccaagggtca ccaccaccac 2520
 caccac 2526

<210> 388
 <211> 842
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 388

Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
 1 5 10 15
 Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu
 20 25 30

Asp Ile His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu
645 650 655

Ala Val Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly
660 665 670

Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile
675 680 685

Pro Tyr Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe
690 695 700

Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys
705 710 715 720

Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp
725 730 735

Leu Asn Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala
740 745 750

Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala
755 760 765

Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu
770 775 780

Leu Gln Val Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala
785 790 795 800

Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro
805 810 815

Leu Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu
820 825 830

Ser Ala Lys Gly His His His His His His
835 840

<210> 389
<211> 42
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 389
gaaaacctcc tcaagcacct ggaccaggta aagccagaaa ac

42

<210> 390
<211> 42
<212> DNA
<213> Artificial

<220>

<223> Synthetic

<400> 390

gttttctggc ttacctggt ccaggtgctt gaggaggttt tc

42

<210> 391

<211> 2526

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 391

| | |
|---|------|
| atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac | 60 |
| ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc | 120 |
| gaaccggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac | 180 |
| gggtacaagg ccgtcttcgt ggtctttgac gccaaaggccc cctccttcg ccacgaggcc | 240 |
| tacgaggcct acaaggcggg gagggccccg acccccgagg acttccccg gcagctcgcc | 300 |
| ctcatcaagg agctggtgga cctcctgggg ttaccgcgc tcgaggtccc cggctacgag | 360 |
| gcggacgacg ttctcgccac cctggccaag aaggcggaaa aggaggggta cgaggtgcgc | 420 |
| atcctcaccg ccgaccgga cctctacca ctctctccg accgcgtcgc cgtcctccac | 480 |
| cccgagggcc acctcatcac ccgggagtgg ctttgggaga agtacggcct caggccggag | 540 |
| cagtgggtgg acttcgcgc cctcgtgggg gaccctccg acaacctccc cggggtcaag | 600 |
| ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc | 660 |
| ctcaagcacc tggaccaggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg | 720 |
| gaagacctca ggctctcctt ggagctctcc cgggtgcga ccgacctccc cctggagggtg | 780 |
| gacctcgccc aggggcggga gcccgaccgg gaggggctta gggccttcct ggagaggctg | 840 |
| gagttcgga gcctcctcca cgagttcggc ctctggagg ccccgcccc cctggaggag | 900 |
| gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctccg ccccgagccc | 960 |
| atgtggggcg agcttaaagc cctggccgcc tgcaggggag gccgcgtgca ccgggcagca | 1020 |
| gacctcttgg cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc | 1080 |
| gccgtcttgg cctcgaggga ggggctagac ctctgcccc gggacgacct catgctctc | 1140 |
| gcctacctcc tggaccttc gaacaccacc cccgaggggg tggcgcgcg ctacgggggg | 1200 |
| gagtggacgg aggacgcgc ccaccgggcc ctctctcgg agaggctcca tcggaacctc | 1260 |

| | |
|---|------|
| cttaagcgcc tcgaggggga ggagaagctc ctttggctct accacgaggt ggaaaagccc | 1320 |
| ctctcccggg tcctggccca tatggaggcc accggggtac ggcgggacgt ggcctacctt | 1380 |
| caggcccttt ccctggagct tgcggaggag atccgccgcc tcgaggagga ggtcttccgc | 1440 |
| ttggcgggcc accccttcaa cctcaactcc cgggaccagc tggaaagggg gctctttgac | 1500 |
| gagcttaggc ttccgcctt gaagaagacg aagaagacag gcaagcgctc caccagcgcc | 1560 |
| gcggtgctgg aggccctacg ggaggccac cccatcgtgg agaagatcct ccagcaccgg | 1620 |
| gagctcacca agctcaagaa cacctacgtg gacccctcc caagcctcgt ccacccgagg | 1680 |
| acgggccgcc tccacaccg cttcaaccag acggccacgg ccacggggag gcttagtagc | 1740 |
| tccgaccca acctgcagaa catccccgtc cgcacccct tgggccagag gatccgccgg | 1800 |
| gccttcgtgg ccgaggcggg ttgggcgttg gtggccctgg actatagcca gatagagctc | 1860 |
| cgctcctcg cccacctctc cggggacgaa aacctgatca gggcttcca ggaggggaag | 1920 |
| gacatccaca ccagaccgc aagctggatg ttcggcgctc ccccgaggc cgtggacccc | 1980 |
| ctgatcgccc gggcgccaa gacggtgaac ttcggcgctc tctacggcat gtccgccc | 2040 |
| aggctctccc aggagcttgc catccctac gaggaggcgg tggcctttat agagcgctac | 2100 |
| ttcaaagct tcccaaggt gcgggcctgg atagaaaaga ccctggagga ggggaggaag | 2160 |
| cggggctacg tggaaaccct cttcggaaga aggcgctacg tgcccgacct caacgcccgg | 2220 |
| gtgaagagcg tcaggaggc cgcgagcgc atggccttca acatgcccg ccagggcacc | 2280 |
| gccgccgacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg | 2340 |
| gcccgcagtc tcctccaggt cgccaacgag ctctctctgg agggcccca agcgcgggcc | 2400 |
| gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatccct cgccgtgccc | 2460 |
| ctggaggtgg aggtgggat gggggaggac tggctttccg ccaagggtca ccaccaccac | 2520 |
| caccac | 2526 |

<210> 392
 <211> 842
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 392

Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val

| | | | |
|-------------|---------------------|---------------------|-------------------------|
| 1 | 5 | 10 | 15 |
| Leu Leu Val | Asp Gly His His Leu | Ala Tyr Arg Thr Phe | Phe Ala Leu |
| | 20 | 25 | 30 |
| Lys Gly Leu | Thr Thr Ser Arg Gly | Glu Pro Val Gln | Ala Val Tyr Gly |
| | 35 | 40 | 45 |
| Phe Ala Lys | Ser Leu Leu Lys | Ala Leu Lys | Glu Asp Gly Tyr Lys Ala |
| | 50 | 55 | 60 |
| Val Phe Val | Val Phe Asp Ala Lys | Ala Pro Ser Phe | Arg His Glu Ala |
| 65 | 70 | 75 | 80 |
| Tyr Glu Ala | Tyr Lys Ala Gly | Arg Ala Pro Thr | Pro Glu Asp Phe Pro |
| | 85 | 90 | 95 |
| Arg Gln Leu | Ala Leu Ile Lys | Glu Leu Val Asp | Leu Leu Gly Phe Thr |
| | 100 | 105 | 110 |
| Arg Leu Glu | Val Pro Gly Tyr | Glu Ala Asp Asp | Val Leu Ala Thr Leu |
| | 115 | 120 | 125 |
| Ala Lys Lys | Ala Glu Lys Glu | Gly Tyr Glu Val | Arg Ile Leu Thr Ala |
| | 130 | 135 | 140 |
| Asp Arg Asp | Leu Tyr Gln Leu | Val Ser Asp Arg | Val Ala Val Leu His |
| 145 | 150 | 155 | 160 |
| Pro Glu Gly | His Leu Ile Thr | Pro Glu Trp Leu | Trp Glu Lys Tyr Gly |
| | 165 | 170 | 175 |
| Leu Arg Pro | Glu Gln Trp Val | Asp Phe Arg Ala | Leu Val Gly Asp Pro |
| | 180 | 185 | 190 |
| Ser Asp Asn | Leu Pro Gly Val | Lys Gly Ile Gly | Glu Lys Thr Ala Leu |
| | 195 | 200 | 205 |
| Lys Leu Leu | Lys Glu Trp Gly | Ser Leu Glu Asn | Leu Leu Lys His Leu |
| | 210 | 215 | 220 |
| Asp Gln Val | Lys Pro Glu Asn | Val Arg Glu Lys | Ile Lys Ala His Leu |
| 225 | 230 | 235 | 240 |
| Glu Asp Leu | Arg Leu Ser Leu | Glu Leu Ser Arg | Val Arg Thr Asp Leu |
| | 245 | 250 | 255 |
| Pro Leu Glu | Val Asp Leu Ala | Gln Gly Arg Glu | Pro Asp Arg Glu Gly |
| | 260 | 265 | 270 |
| Leu Arg Ala | Phe Leu Glu Arg | Leu Glu Phe Gly | Ser Leu Leu His Glu |
| | 275 | 280 | 285 |
| Phe Gly Leu | Leu Glu Ala Pro | Ala Pro Leu Glu | Glu Ala Pro Trp Pro |
| | 290 | 295 | 300 |
| Pro Pro Glu | Gly Ala Phe Val | Gly Phe Val Leu | Ser Arg Pro Glu Pro |

| 610 | 615 | 620 |
|--|-----|-----|
| His Leu Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys 625 630 635 640 | | |
| Asp Ile His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu 645 650 655 | | |
| Ala Val Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly 660 665 670 | | |
| Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile 675 680 685 | | |
| Pro Tyr Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe 690 695 700 | | |
| Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys 705 710 715 720 | | |
| Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp 725 730 735 | | |
| Leu Asn Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala 740 745 750 | | |
| Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala 755 760 765 | | |
| Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu 770 775 780 | | |
| Leu Gln Val Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala 785 790 795 800 | | |
| Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro 805 810 815 | | |
| Leu Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu 820 825 830 | | |
| Ser Ala Lys Gly His His His His His His 835 840 | | |

<210> 393
 <211> 33
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 393
 gagctctccc ggggtgcacac cgacctcccc ctg

<210> 394

<211> 33
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 394
 cagggggagg tcggtgtgca cccgggagag etc

33

<210> 395
 <211> 2526
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 395
 atgaattcgc aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
 ggccaccacc tggectaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
 gaaccggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac 180
 ggggtacaagg ccgtcttcgt ggtctttgac gccaaaggccc cctccttcgc ccacgaggcc 240
 tacgaggcct acaaggcggg gagggccccg acccccagag acttccccgc gcagctcgcc 300
 ctcatcaagg agctggtgga cctcctgggg tttaccgcgc tcgagggtccc cggtacgag 360
 gcggacgacg ttctcgccac cctggccaag aaggcggaag aggaggggta cgaggtgcgc 420
 atcctcaccg ccgaccgcga cctctaccaa ctgctctccg accgcgtcgc cgtcctccac 480
 cccgaggggc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag 540
 cagtgggtgg acttcgcgc cctcgtgggg gaccctccg acaacctccc cggggtcaag 600
 ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc 660
 ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg 720
 gaagacctca ggctctcctt ggagctctcc cgggtgcaca ccgacctccc cctggagggtg 780
 gacctcgccc aggggcggga gcccgaccgg gaggggctta gggccttcct ggagaggctg 840
 gagttcggca gcctcctcca cgagttcggc ctctggagg cccccgccc cctggaggag 900
 gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctccc ccccgagccc 960
 atgtgggcgg agcttaaagc cctggccgcc tgcaggggcg gccgcgtgca ccgggcagca 1020
 gaccctcttg cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc 1080
 gccgtcttg cctcgaggga ggggctagac ctggtgccc gggacgacct catgctctc 1140

gcctacctcc tggacccttc gaacaccacc cccgaggggg tggcgcgggc ctacgggggg 1200
gagtggacgg aggacgccgc ccaccgggcc ctctctctcg agaggctcca tcggaacctc 1260
cttaagcgcc tcgaggggga ggagaagctc ctttggtctt accacgaggt ggaaaagccc 1320
ctctcccggg tcctggccca tatggaggcc accgggggtac ggcgggacgt ggcctacctt 1380
caggcccttt ccttggagct tgcggaggag atccgccgcc tcgaggagga ggtcttccgc 1440
ttggcgggcc accccttcaa cctcaactcc cgggaccagc tggaaagggt gctctttgac 1500
gagcttaggc ttccgcctt gaagaagacg aagaagacag gcaagcgctc caccagcgcc 1560
gcggtgctgg aggccctacg ggaggccac cccatcgtag agaagatcct ccagcaccgg 1620
gagctcacca agctcaagaa cacctacgtg gacccctcc caagcctcgt ccacccgagg 1680
acgggcccgc tccacacccg cttcaaccag acggccacgg ccacggggag gcttagtagc 1740
tccgacccca acctgcagaa catccccgtc cgcacccctt tgggccagag gatccgccgg 1800
gccttcgtgg ccgaggcggg ttgggcgttg gtggccctgg actatagcca gatagagctc 1860
cgcgtcctcg cccacctctc cggggacgaa aacctgatca gggctctcca ggaggggaag 1920
gacatccaca cccagaccgc aagctggatg ttcggcgctc ccccgaggc cgtggacccc 1980
ctgatgcgcc gggcgggcaa gacggtgaac ttcggcgctc tctacggcat gtccgcccac 2040
aggctctccc aggagcttgc catcccctac gaggaggcgg tggcctttat agagcgctac 2100
ttccaaagct tccccaaggt gcgggcctgg atagaaaaga ccctggagga ggggaggaag 2160
cggggctacg tggaaaccct cttcggaaga aggcgctacg tgcccgacct caacgcccgg 2220
gtgaagagcg tcagggaggc cgcggagcgc atggccttca acatgcccgt ccagggcacc 2280
gccgccgacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg 2340
gcccgcacgc tcctccaggt cgccaacgag ctctctctgg agggccccca agcgcgggcc 2400
gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatcccct cgccgtgccc 2460
ctggaggtgg aggtggggat gggggaggac tggttttccg ccaagggtca ccaccaccac 2520
caccac 2526

<210> 396
<211> 842
<212> PRT
<213> Artificial

<220>
<223> Synthetic

Pro Leu Gly Gln Arg Ile Arg Arg Ala Phe Val Ala Glu Ala Gly Trp
595 600 605

Ala Leu Val Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala
610 615 620

His Leu Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys
625 630 635 640

Asp Ile His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu
645 650 655

Ala Val Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly
660 665 670

Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile
675 680 685

Pro Tyr Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe
690 695 700

Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys
705 710 715 720

Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp
725 730 735

Leu Asn Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala
740 745 750

Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala
755 760 765

Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu
770 775 780

Leu Gln Val Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala
785 790 795 800

Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro
805 810 815

Leu Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu
820 825 830

Ser Ala Lys Gly His His His His His His
835 840

<210> 397
<211> 33
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 397

gagctctccc ggggtgcacac cgacctctc ctg 33

<210> 398
<211> 33
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 398
caggaggagg tcggtgtgca cccgggagag ctc 33

<210> 399
<211> 2526
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 399
atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
gaaccggtgc aggcggtcta cggtctcgcc aagagcctcc tcaaggccct gaaggaggac 180
gggtacaagg ccgtcttcgt ggtctttgac gccaaaggccc cctccttcctg ccacgaggcc 240
tacgaggcct acaaggcggg gagggccccc acccccgagg acttcccccg gcagctcgcc 300
ctcatcaagg agctggtgga cctcctgggg ttaccgcgc tcgaggtccc cggctacgag 360
gcggacgacg ttctcgccac cctggccaag aaggcggaaa aggaggggta cgagggtgcgc 420
atcctcaccg ccgaccgga cctctaccaa ctgctctccg accgcgtcgc cgtcctccac 480
cccgagggcc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag 540
cagtgggtgg acttccgcgc cctcgtgggg gaccctccg acaacctccc cggggtcaag 600
ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc 660
ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggccacctg 720
gaagacctca ggctctcctt ggagctctcc cgggtgcaca ccgacctct cctggagggtg 780
gacctcgccc aggggcggga gcccgaccgg gaggggctta gggccttcct ggagaggctg 840
gagttcgga gcctcctcca cgagttcggc ctctggagg ccccgcccc cctggaggag 900
gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctcccg ccccgagccc 960
atgtgggcgg agcttaaagc cctggccgcc tgcaggggcg gccgcgtgca ccgggcagca 1020

| | | | | | | |
|------------|------------|------------|-------------|------------|------------|------|
| gaccccttgg | cggggctaaa | ggacctcaag | gaggtccggg | gcctcctcgc | caaggacctc | 1080 |
| gccgtcttgg | cctcgaggga | ggggctagac | ctcgtgcccg | gggacgaccc | catgctcctc | 1140 |
| gcctacctcc | tggaccttcc | gaacaccacc | cccgaggggg | tggcgcgggc | ctacgggggg | 1200 |
| gagtggacgg | aggacgccgc | ccaccggggc | ctcctctcgg | agaggctcca | tcggaacctc | 1260 |
| cttaagcgcc | tcgaggggga | ggagaagctc | ctttggctct | accacgaggt | ggaaaagccc | 1320 |
| ctctcccggg | tcctggccca | tatggaggcc | accgggggtac | ggcgggacgt | ggcctacctt | 1380 |
| caggcccttt | ccctggagct | tgcgaggag | atccgccgcc | tcgaggagga | ggtcttccgc | 1440 |
| ttggcggggc | accccttcaa | cctcaactcc | cgggaccagc | tggaaagggg | gctctttgac | 1500 |
| gagcttaggc | ttcccgctt | gaagaagacg | aagaagacag | gcaagcgctc | caccagcgcc | 1560 |
| gcggtgctgg | aggccctacg | ggaggcccac | cccatcgtag | agaagatcct | ccagcaccgg | 1620 |
| gagctcacca | agctcaagaa | cacctacgtg | gacccctcct | caagcctcgt | ccacccgagg | 1680 |
| acgggcccgc | tccacaccgc | cttcaaccag | acggccacgg | ccacggggag | gcttagtagc | 1740 |
| tccgacccca | acctgcagaa | catccccgtc | cgcacccctt | tgggccagag | gatccgcccg | 1800 |
| gccttcgtgg | ccgaggcggg | ttgggcgttg | gtggccctgg | actatagcca | gatagagctc | 1860 |
| cgcgtcctcg | cccacctctc | cggggacgaa | aacctgatca | gggtcttcca | ggaggggaag | 1920 |
| gacatccaca | cccagaccgc | aagctggatg | ttcggcgctc | ccccggaggc | cgtggacccc | 1980 |
| ctgatgcgcc | gggcggccaa | gacggtgaac | ttcggcgctc | tctacggcat | gtccgcccac | 2040 |
| aggctctccc | aggagcttgc | catcccctac | gaggaggcgg | tggcctttat | agagcgctac | 2100 |
| ttccaaagct | tccccaaggt | gcgggccttg | atagaaaaga | ccctggagga | ggggaggaag | 2160 |
| cggggctacg | tggaaaccct | cttcggaaga | aggcgctacg | tgcccacact | caacgcccgg | 2220 |
| gtgaagagcg | tcagggaggc | cgcggagcgc | atggccttca | acatgcccgt | ccagggcacc | 2280 |
| gccgccgacc | tcatgaagct | cgccatggtg | aagctcttcc | ccgcctccg | ggagatgggg | 2340 |
| gcccgcacgc | tcctccaggt | cgccaacgag | ctcctcctgg | aggcccccca | agcgcgggcc | 2400 |
| gaggaggtgg | cggctttggc | caaggaggcc | atggagaagg | cctatcccct | cgcgctgccc | 2460 |
| ctggaggtgg | aggtggggat | gggggaggac | tggctttccg | ccaagggtca | ccaccaccac | 2520 |
| caccac | | | | | | 2526 |

<210> 400
 <211> 842
 <212> PRT

<213> Artificial

<220>

<223> Synthetic

<400> 400

Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
1 5 10 15
Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu
20 25 30
Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
35 40 45
Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala
50 55 60
Val Phe Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala
65 70 75 80
Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro
85 90 95
Arg Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr
100 105 110
Arg Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu
115 120 125
Ala Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala
130 135 140
Asp Arg Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His
145 150 155 160
Pro Glu Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly
165 170 175
Leu Arg Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro
180 185 190
Ser Asp Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu
195 200 205
Lys Leu Leu Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu
210 215 220
Asp Arg Val Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu
225 230 235 240
Glu Asp Leu Arg Leu Ser Leu Glu Leu Ser Arg Val His Thr Asp Leu
245 250 255
Leu Leu Glu Val Asp Leu Ala Gln Gly Arg Glu Pro Asp Arg Glu Gly
260 265 270

<220>

<223> Synthetic

<400> 401

atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
gaaccggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac 180
gggtacaagg ccgtcttcgt ggtctttgac gccaaggccc cctccttcgc ccacgaggcc 240
tacgaggcct acaaggcggg gagggcccg acccccgagg acttccccgc gcagctcgcc 300
ctcatcaagg agctggtgga cctcctgggg tttaccgcgc tcgagggtccc cggctacgag 360
gcggaacgacg ttctcgccac cctggccaag aaggcggaaa aggaggggta cgaggtgctc 420
atcctcaccg ccgaccgcga cctctaccaa ctgctctccg accgcgtcgc cgtcctccac 480
cccgagggcc acctcatcac ccgggagtgg ctttgggaga agtacggcct caggccggag 540
cagtgggtgg acttccgcgc cctcgtgggg gacccctccg acaacctccc cggggtcaag 600
ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc 660
ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg 720
gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctccc cctggagggtg 780
gacctcgccc aggggcggga gcccgaccgg gaggggctta gggccttcct ggagaggctg 840
gagttcgga gcctcctcca cgagttcggc ctcttgagg cccccgccc cctggaggag 900
gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctcccg ccccgagccc 960
atgtgggcgg agcttaaagc cctggccgcc tgcaggggcg gccgcgtgca ccgggcagca 1020
gaccccttgg cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc 1080
gccgtcttgg cctcgaggga ggggctagac ctggtgccc gggacgacct catgctcctc 1140
gcctacctcc tggaccttc gaacaccacc cccgaggggg tggcgcgcg ctacgggggg 1200
gagtggacgg aggacgcgc ccaccgggccc ctctctcgg agaggctcca tcggaacctc 1260
cttaagcgcc tcgaggggga ggagaagctc gtttggctct accacgaggt ggaaaagccc 1320
ctctcccggg tcctggccca tatggaggcc accggggtac ggcgggacgt ggcctacctt 1380
caggcccttt ccttgagct tgcggaggag atccgccgcc tcgaggagga ggtcttccgc 1440
ttggcggggc accccttcaa cctcaactcc cgggaccagc tggaaagggt gctctttgac 1500
gagcttaggc tccccgctt gaagaagacg aagaagacag gcaagcgctc caccagcgcc 1560
gcggtgctgg aggcctacg ggaggccac cccatcgtgg agaagatcct ccagcaccgg 1620

CG43436.1

Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro
 85 90 95
 Arg Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr
 100 105 110
 Arg Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu
 115 120 125
 Ala Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala
 130 135 140
 Asp Arg Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His
 145 150 155 160
 Pro Glu Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly
 165 170 175
 Leu Arg Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro
 180 185 190
 Ser Asp Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu
 195 200 205
 Lys Leu Leu Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu
 210 215 220
 Asp Arg Val Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu
 225 230 235 240
 Glu Asp Leu Arg Leu Ser Leu Glu Leu Ser Arg Val Arg Thr Asp Leu
 245 250 255
 Pro Leu Glu Val Asp Leu Ala Gln Gly Arg Glu Pro Asp Arg Glu Gly
 260 265 270
 Leu Arg Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu
 275 280 285
 Phe Gly Leu Leu Glu Ala Pro Ala Pro Leu Glu Glu Ala Pro Trp Pro
 290 295 300
 Pro Pro Glu Gly Ala Phe Val Gly Phe Val Leu Ser Arg Pro Glu Pro
 305 310 315 320
 Met Trp Ala Glu Leu Lys Ala Leu Ala Ala Cys Arg Gly Gly Arg Val
 325 330 335
 His Arg Ala Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val
 340 345 350
 Arg Gly Leu Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly
 355 360 365
 Leu Asp Leu Val Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu
 370 375 380

Pro Tyr Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe
690 695 700

Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys
705 710 715 720

Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp
725 730 735

Leu Asn Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala
740 745 750

Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala
755 760 765

Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu
770 775 780

Leu Gln Val Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala
785 790 795 800

Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro
805 810 815

Leu Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu
820 825 830

Ser Ala Lys Gly
835

<210> 403
<211> 66
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<220>
<221> modified_base
<222> (19)..(51)
<223> The bases in these positions are a synthesis of 91% base shown and 3% all other bases

<400> 403
ctccatcgga acctccttaa gcgcctcgag ggggaggaga agctcctttg gctctaccac 60

gaggtg 66

<210> 404
<211> 18
<212> DNA
<213> Artificial

<220>

| 20 | | | | | 25 | | | | | 30 | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| Lys | Gly | Leu | Thr | Thr | Ser | Arg | Gly | Glu | Pro | Val | Gln | Ala | Val | Tyr | Gly | |
| 35 | | | | | 40 | | | | | 45 | | | | | | |
| Phe | Ala | Lys | Ser | Leu | Leu | Lys | Ala | Leu | Lys | Glu | Asp | Gly | Tyr | Lys | Ala | |
| 50 | | | | | 55 | | | | | 60 | | | | | | |
| Val | Phe | Val | Val | Phe | Asp | Ala | Lys | Ala | Pro | Ser | Phe | Arg | His | Glu | Ala | |
| 65 | | | | | 70 | | | | | 75 | | | | | 80 | |
| Tyr | Glu | Ala | Tyr | Lys | Ala | Gly | Arg | Ala | Pro | Thr | Pro | Glu | Asp | Phe | Pro | |
| 85 | | | | | 90 | | | | | 95 | | | | | | |
| Arg | Gln | Leu | Ala | Leu | Ile | Lys | Glu | Leu | Val | Asp | Leu | Leu | Gly | Phe | Thr | |
| 100 | | | | | 105 | | | | | 110 | | | | | | |
| Arg | Leu | Glu | Val | Pro | Gly | Tyr | Glu | Ala | Asp | Asp | Val | Leu | Ala | Thr | Leu | |
| 115 | | | | | 120 | | | | | 125 | | | | | | |
| Ala | Lys | Lys | Ala | Glu | Lys | Glu | Gly | Tyr | Glu | Val | Arg | Ile | Leu | Thr | Ala | |
| 130 | | | | | 135 | | | | | 140 | | | | | | |
| Asp | Arg | Asp | Leu | Tyr | Gln | Leu | Val | Ser | Asp | Arg | Val | Ala | Val | Leu | His | |
| 145 | | | | | 150 | | | | | 155 | | | | | 160 | |
| Pro | Glu | Gly | His | Leu | Ile | Thr | Pro | Glu | Trp | Leu | Trp | Glu | Lys | Tyr | Gly | |
| 165 | | | | | 170 | | | | | 175 | | | | | | |
| Leu | Arg | Pro | Glu | Gln | Trp | Val | Asp | Phe | Arg | Ala | Leu | Val | Gly | Asp | Pro | |
| 180 | | | | | 185 | | | | | 190 | | | | | | |
| Ser | Asp | Asn | Leu | Pro | Gly | Val | Lys | Gly | Ile | Gly | Glu | Lys | Thr | Ala | Leu | |
| 195 | | | | | 200 | | | | | 205 | | | | | | |
| Lys | Leu | Leu | Lys | Glu | Trp | Gly | Ser | Leu | Glu | Asn | Leu | Leu | Lys | Asn | Leu | |
| 210 | | | | | 215 | | | | | 220 | | | | | | |
| Asp | Arg | Val | Lys | Pro | Glu | Asn | Val | Arg | Glu | Lys | Ile | Lys | Ala | His | Leu | |
| 225 | | | | | 230 | | | | | 235 | | | | | 240 | |
| Glu | Asp | Leu | Arg | Leu | Ser | Leu | Glu | Leu | Ser | Arg | Val | Arg | Thr | Asp | Leu | |
| 245 | | | | | 250 | | | | | 255 | | | | | | |
| Pro | Leu | Glu | Val | Asp | Leu | Ala | Gln | Gly | Arg | Glu | Pro | Asp | Arg | Glu | Gly | |
| 260 | | | | | 265 | | | | | 270 | | | | | | |
| Leu | Arg | Ala | Phe | Leu | Glu | Arg | Leu | Glu | Phe | Gly | Ser | Leu | Leu | His | Glu | |
| 275 | | | | | 280 | | | | | 285 | | | | | | |
| Phe | Gly | Leu | Leu | Glu | Ala | Pro | Ala | Pro | Leu | Glu | Glu | Ala | Pro | Trp | Pro | |
| 290 | | | | | 295 | | | | | 300 | | | | | | |
| Pro | Pro | Glu | Gly | Ala | Phe | Val | Gly | Phe | Val | Leu | Ser | Arg | Pro | Glu | Pro | |
| 305 | | | | | 310 | | | | | 315 | | | | | 320 | |
| Met | Trp | Ala | Glu | Leu | Lys | Ala | Leu | Ala | Ala | Cys | Arg | Gly | Gly | Arg | Val | |

| | | |
|---|-----|-----|
| 325 | 330 | 335 |
| His Arg Ala Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val | | |
| 340 | 345 | 350 |
| Arg Gly Leu Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly | | |
| 355 | 360 | 365 |
| Leu Asp Leu Val Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu | | |
| 370 | 375 | 380 |
| Asp Pro Ser Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly | | |
| 385 | 390 | 395 |
| Glu Trp Thr Glu Asp Ala Ala His Arg Ala Leu Leu Ser Glu Arg Leu | | |
| 405 | 410 | 415 |
| His Arg Asn Leu Leu Lys Arg Leu Glu Gly Val Glu Lys Leu Leu Trp | | |
| 420 | 425 | 430 |
| Leu Tyr His Glu Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met | | |
| 435 | 440 | 445 |
| Glu Ala Thr Gly Val Arg Arg Asp Val Ala Tyr Leu Gln Ala Leu Ser | | |
| 450 | 455 | 460 |
| Leu Glu Leu Ala Glu Glu Ile Arg Arg Leu Glu Glu Glu Val Phe Arg | | |
| 465 | 470 | 475 |
| Leu Ala Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg | | |
| 485 | 490 | 495 |
| Val Leu Phe Asp Glu Leu Arg Leu Pro Ala Leu Lys Lys Thr Lys Lys | | |
| 500 | 505 | 510 |
| Thr Gly Lys Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu | | |
| 515 | 520 | 525 |
| Ala His Pro Ile Val Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys | | |
| 530 | 535 | 540 |
| Leu Lys Asn Thr Tyr Val Asp Pro Leu Pro Ser Leu Val His Pro Arg | | |
| 545 | 550 | 555 |
| Thr Gly Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly | | |
| 565 | 570 | 575 |
| Arg Leu Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr | | |
| 580 | 585 | 590 |
| Pro Leu Gly Gln Arg Ile Arg Arg Ala Phe Val Ala Glu Ala Gly Trp | | |
| 595 | 600 | 605 |
| Ala Leu Val Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala | | |
| 610 | 615 | 620 |
| His Leu Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys | | |

625 630 635 640

Asp Ile His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu
645 650 655

Ala Val Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly
660 665 670

Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile
675 680 685

Pro Tyr Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe
690 695 700

Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys
705 710 715 720

Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp
725 730 735

Leu Asn Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala
740 745 750

Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala
755 760 765

Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu
770 775 780

Leu Gln Val Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala
785 790 795 800

Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro
805 810 815

Leu Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu
820 825 830

Ser Ala Lys Gly
835

<210> 407
<211> 2508
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 407
atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
ggccaccacc tggcctaccg caccttcttc gcctgaagg gcctcaccac gagccggggc 120
gaaccggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac 180
gggtacaagg ccgtcttcgt ggtctttgac gccaaagccc cctccttcgc ccacgaggcc 240

gacatccaca cccagaccgc aagctggatg ttcggcgctcc ccccgagggc cgtggacccc 1980
 ctgatgcgcc gggcggccaa gacggtgaac ttcggcgctcc tctacggcat gtccgcccac 2040
 aggctctccc aggagcttgc catcccctac gaggaggcgg tggcctttat agagcgctac 2100
 ttccaaagct tccccaaggt gcgggcctgg atagaaaaga ccttgaggga ggggaggaag 2160
 cggggctacg tggaaaccct ctccggaaga aggcgctacg tgcccgcact caacgcccgg 2220
 gtgaagagcg tcaggagggc cgcggagcgc atggccttca acatgcccggt ccagggcacc 2280
 gccgccgacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg 2340
 gcccgcatgc tcctccaggt cgccaacgag ctctcctgg aggccccca agcgcggggc 2400
 gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatcccct cgcggtgccc 2460
 ctggaggtgg aggtggggat gggggaggac tggctttccg ccaagggt 2508

<210> 408
 <211> 836
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 408

Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
 1 5 10 15
 Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu
 20 25 30
 Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
 35 40 45
 Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala
 50 55 60
 Val Phe Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala
 65 70 75 80
 Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro
 85 90 95
 Arg Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr
 100 105 110
 Arg Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu
 115 120 125
 Ala Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala
 130 135 140

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asp | Arg | Asp | Leu | Tyr | Gln | Leu | Val | Ser | Asp | Arg | Val | Ala | Val | Leu | His | 145 | 150 | 155 | 160 |
| Pro | Glu | Gly | His | Leu | Ile | Thr | Pro | Glu | Trp | Leu | Trp | Glu | Lys | Tyr | Gly | 165 | 170 | 175 | |
| Leu | Arg | Pro | Glu | Gln | Trp | Val | Asp | Phe | Arg | Ala | Leu | Val | Gly | Asp | Pro | 180 | 185 | 190 | |
| Ser | Asp | Asn | Leu | Pro | Gly | Val | Lys | Gly | Ile | Gly | Glu | Lys | Thr | Ala | Leu | 195 | 200 | 205 | |
| Lys | Leu | Leu | Lys | Glu | Trp | Gly | Ser | Leu | Glu | Asn | Leu | Leu | Lys | Asn | Leu | 210 | 215 | 220 | |
| Asp | Arg | Val | Lys | Pro | Glu | Asn | Val | Arg | Glu | Lys | Ile | Lys | Ala | His | Leu | 225 | 230 | 235 | 240 |
| Glu | Asp | Leu | Arg | Leu | Ser | Leu | Glu | Leu | Ser | Arg | Val | Arg | Thr | Asp | Leu | 245 | 250 | 255 | |
| Pro | Leu | Glu | Val | Asp | Leu | Ala | Gln | Gly | Arg | Glu | Pro | Asp | Arg | Glu | Gly | 260 | 265 | 270 | |
| Leu | Arg | Ala | Phe | Leu | Glu | Arg | Leu | Glu | Phe | Gly | Ser | Leu | Leu | His | Glu | 275 | 280 | 285 | |
| Phe | Gly | Leu | Leu | Glu | Ala | Pro | Ala | Pro | Leu | Glu | Glu | Ala | Pro | Trp | Pro | 290 | 295 | 300 | |
| Pro | Pro | Glu | Gly | Ala | Phe | Val | Gly | Phe | Val | Leu | Ser | Arg | Pro | Glu | Pro | 305 | 310 | 315 | 320 |
| Met | Trp | Ala | Glu | Leu | Lys | Ala | Leu | Ala | Ala | Cys | Arg | Gly | Gly | Arg | Val | 325 | 330 | 335 | |
| His | Arg | Ala | Ala | Asp | Pro | Leu | Ala | Gly | Leu | Lys | Asp | Leu | Lys | Glu | Val | 340 | 345 | 350 | |
| Arg | Gly | Leu | Leu | Ala | Lys | Asp | Leu | Ala | Val | Leu | Ala | Ser | Arg | Glu | Gly | 355 | 360 | 365 | |
| Leu | Asp | Leu | Val | Pro | Gly | Asp | Asp | Pro | Met | Leu | Leu | Ala | Tyr | Leu | Leu | 370 | 375 | 380 | |
| Asp | Pro | Ser | Asn | Thr | Thr | Pro | Glu | Gly | Val | Ala | Arg | Arg | Tyr | Gly | Gly | 385 | 390 | 395 | 400 |
| Glu | Trp | Thr | Glu | Asp | Ala | Ala | His | Arg | Ala | Leu | Leu | Ser | Glu | Arg | Leu | 405 | 410 | 415 | |
| His | Arg | Asn | Leu | Leu | Lys | Arg | Asn | Glu | Gly | Lys | Glu | Lys | Leu | Leu | Trp | 420 | 425 | 430 | |
| Leu | Tyr | His | Glu | Val | Glu | Lys | Pro | Leu | Ser | Arg | Val | Leu | Ala | His | Met | 435 | 440 | 445 | |

gccccctggc ccccgccgga aggggccttc gtgggcttcg tctctctccg ccccgagccc 960
 atgtggggcg agcttaaagc cctggccgcc tgcaggggcg gccgcgtgca ccgggcagca 1020
 gacccttgg cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc 1080
 gccgtcttgg cctcgaggga ggggctagac ctctgtcccc gggacgaccc catgctcctc 1140
 gcctacctcc tggaccttc gaacaccacc cccgaggggg tggcgcgggc ctacgggggg 1200
 gagtggacgg aggacgccgc ccaccgggccc ctctctcgg agaggctcca tcggaacctc 1260
 cttaagcgct tcgaggggga ggagaagctc ctttgcctct accacgaggt ggaaaagccc 1320
 ctctcccggg tcttgcccc a tatggaggcc accggggtac ggcgggacgt ggcctacctt 1380
 caggcccttt ccctggagct tgcggaggag atccgccgcc tcgaggagga ggtcttccgc 1440
 ttggcgggcc accccttcaa cctcaactcc cgggaccagc tggaaagggt gctctttgac 1500
 gagcttaggc ttccgcctt gaagaagacg aagaagacag gcaagcgctc caccagcgcc 1560
 gcggtgctgg aggccctacg ggaggccac cccatcgtag agaagatcct ccagcaccgg 1620
 gagctcacca agctcaagaa cacctacgtg gacccctcc caagcctcgt ccaccgagg 1680
 acgggcccgc tccacaccgc cttcaaccag acggccacgg ccacggggag gcttagtagc 1740
 tccgacccca acctgcagaa catccccgtc cgcacccctc tgggcccagag gatccgccgg 1800
 gccttcgtgg ccgaggcggg ttgggcttg gtggccctgg actatagcca gatagagctc 1860
 cgcgtcctcg ccacctctc cggggacgaa aacctgatca gggctctcca ggaggggaag 1920
 gacatccaca ccagaccgc aagctggatg ttcggcgtcc ccccgagggc cgtggacccc 1980
 ctgatgcgcc gggcgccaa gacggtgaac ttcggcgtcc tctacggcat gtccgcccat 2040
 aggctctccc aggagcttgc catcccctac gaggaggcgg tggcctttat agagcgctac 2100
 ttcaaagct tcccaaggt gcgggcctgg atagaaaaga ccctggagga ggggaggaag 2160
 cggggctacg tggaaaccct cttcggaaga aggcgtacg tgcccacact caacgcccgg 2220
 gtgaagagcg tcagggaggc cgcggagcgc atggccttca acatgcccgt ccagggcacc 2280
 gccgccgacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg 2340
 gccgcacatg tctccaggt cgccaacgag ctctccttgg agggccccca agcgcgggcc 2400
 gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatccctc cgcgtgccc 2460
 ctggaggtgg aggtggggat gggggaggac tggctttccg ccaagggt 2508

<210> 410

Thr Gly Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly
565 570 575

Arg Leu Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr
580 585 590

Pro Leu Gly Gln Arg Ile Arg Arg Ala Phe Val Ala Glu Ala Gly Trp
595 600 605

Ala Leu Val Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala
610 615 620

His Leu Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys
625 630 635 640

Asp Ile His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu
645 650 655

Ala Val Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly
660 665 670

Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile
675 680 685

Pro Tyr Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe
690 695 700

Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys
705 710 715 720

Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp
725 730 735

Leu Asn Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala
740 745 750

Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala
755 760 765

Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu
770 775 780

Leu Gln Val Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala
785 790 795 800

Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro
805 810 815

Leu Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu
820 825 830

Ser Ala Lys Gly
835

<210> 411
<211> 2508
<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 411

atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
gaaccggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac 180
gggtacaagg ccgtcttcgt ggtctttgac gccaaggccc cctccttcgc ccacgaggcc 240
tacgaggcct acaaggcggg gagggccccg acccccagagg acttccccgc gcagctcgcc 300
ctcatcaagg agctggtgga cctcctgggg tttaccgcgc tcgagggtccc cggctacgag 360
gcgagcagcg ttctcgccac cctggccaag aaggcggaag aggaggggta cgaggtgcgc 420
atcctcaccg ccgaccgcga cctctaccaa ctctctccg accgcgtcgc cgtcctccac 480
cccaggggcc acctcatcac ccgggagtgg ctttgggaga agtacggcct caggccggag 540
cagtgggtgg acttccgcgc cctcgtgggg gaccctccg acaacctccc cggggtcaag 600
ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc 660
ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg 720
gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctccc cctggaggtg 780
gacctcgccc aggggcggga gcccgaccgg gaggggctta gggccttcct ggagaggctg 840
gagttcggca gcctcctcca cgagttcggc ctcttgagg ccccgcccc cctggaggag 900
gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctcccg ccccgagccc 960
atgtgggcgg agcttaaagc cctggccgcc tgcaggggcg gccgcgtgca ccgggcagca 1020
gacccttggt cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc 1080
gccgtcttgg cctcgaggga ggggctagac ctctgcccc gggacgacct catgctcctc 1140
gcctacctcc tggaccttc gaacaccacc cccgaggggg tggcgcggcg ctacgggggg 1200
gagtggacgg aggacgccgc ccaccgggcc ctctctcgg agaggctcca tcggaacctc 1260
cttaagcgcc tcgaggggga ggagaagctc ctttggctct accacgaggt ggaaaagccc 1320
ctctcccggt tcctggccca tatggaggcc accggggtac ggcgggacgt ggcctacctt 1380
caggcccttt ccctggagct tgcggaggag atccgccgcc tcgaggagga ggtcttcgc 1440
ttggcgggcc accccttcaa cctcaactcc cgggaccage tggaaagggt gctctttgac 1500
gagcttaggc ttcccttttt gaagaagacg aagaagacag gcaagcgctc caccagcgcc 1560

gcggtgctgg aggccctacg ggaggccac cccatcgtgg agaagatcct ccagcaccgg 1620
gagctcacca agctcaagaa cacctacgtg gacccccctcc caagcctcgt ccacccgagg 1680
acggggccgcc tccacacccg cttcaaccag acggccacgg ccacggggag gcttagtagc 1740
tccgacccca acctgcagaa catccccgtc cgcacccccct tgggccagag gatccgccgg 1800
gccttcgtgg ccgaggcggg ttgggcgttg gtggcccttg actatagcca gatagagctc 1860
cgcgtcctcg cccacctctc cggggacgaa aacctgatca gggctctcca ggaggggaag 1920
gacatccaca cccagaccgc aagctggatg ttcggcgtcc ccccgagggc cgtggacccc 1980
ctgatgcgcc gggcggccaa gacggtgaac ttcggcgtcc tctacggcat gtccgcccac 2040
aggctctccc aggagcttgc catcccctac gaggaggcgg tggcctttat agagcgctac 2100
ttccaaagct tcccgaaggt gcgggccttg atagaaaaga ccctggagga ggggaggaag 2160
cggggctacg tggaaaccct cttcgaaga aggcgctacg tgcccgaact caacgcccgg 2220
gtgaagagcg tcaggagggc cgcggagcgc atggccttca acatgcccggt ccagggcacc 2280
gccgcccagc tcattgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg 2340
gcccgcatgc tctccaggt cgccaacgag ctctccttg agggccccca agcgcggggc 2400
gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatccccct cgcggtgccc 2460
ctggaggtgg aggtggggat gggggaggac tggctttccg ccaagggt 2508

<210> 412
<211> 836
<212> PRT
<213> Artificial

<220>
<223> Synthetic

<400> 412

Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
1 5 10 15

Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu
20 25 30

Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
35 40 45

Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala
50 55 60

Val Phe Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala

| | | | | | | |
|---|-----|----|--|-----|--|-----|
| 65 | | 70 | | 75 | | 80 |
| Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro | | | | | | |
| | 85 | | | 90 | | 95 |
| Arg Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr | | | | | | |
| | 100 | | | 105 | | 110 |
| Arg Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu | | | | | | |
| | 115 | | | 120 | | 125 |
| Ala Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala | | | | | | |
| | 130 | | | 135 | | 140 |
| Asp Arg Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His | | | | | | |
| | 145 | | | 150 | | 155 |
| Pro Glu Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly | | | | | | |
| | 165 | | | 170 | | 175 |
| Leu Arg Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro | | | | | | |
| | 180 | | | 185 | | 190 |
| Ser Asp Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu | | | | | | |
| | 195 | | | 200 | | 205 |
| Lys Leu Leu Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu | | | | | | |
| | 210 | | | 215 | | 220 |
| Asp Arg Val Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu | | | | | | |
| | 225 | | | 230 | | 235 |
| Glu Asp Leu Arg Leu Ser Leu Glu Leu Ser Arg Val Arg Thr Asp Leu | | | | | | |
| | 245 | | | 250 | | 255 |
| Pro Leu Glu Val Asp Leu Ala Gln Gly Arg Glu Pro Asp Arg Glu Gly | | | | | | |
| | 260 | | | 265 | | 270 |
| Leu Arg Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu | | | | | | |
| | 275 | | | 280 | | 285 |
| Phe Gly Leu Leu Glu Ala Pro Ala Pro Leu Glu Glu Ala Pro Trp Pro | | | | | | |
| | 290 | | | 295 | | 300 |
| Pro Pro Glu Gly Ala Phe Val Gly Phe Val Leu Ser Arg Pro Glu Pro | | | | | | |
| | 305 | | | 310 | | 315 |
| Met Trp Ala Glu Leu Lys Ala Leu Ala Ala Cys Arg Gly Gly Arg Val | | | | | | |
| | 325 | | | 330 | | 335 |
| His Arg Ala Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val | | | | | | |
| | 340 | | | 345 | | 350 |
| Arg Gly Leu Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly | | | | | | |
| | 355 | | | 360 | | 365 |
| Leu Asp Leu Val Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu | | | | | | |

| | | |
|-------------------------|-------------------------|-----------------|
| 675 | 680 | 685 |
| Pro Tyr Glu Glu Ala Val | Ala Phe Ile Glu Arg Tyr | Phe Gln Ser Phe |
| 690 | 695 | 700 |
| Pro Lys Val Arg Ala Trp | Ile Glu Lys Thr Leu Glu | Glu Gly Arg Lys |
| 705 | 710 | 715 |
| Arg Gly Tyr Val Glu Thr | Leu Phe Gly Arg Arg Arg | Tyr Val Pro Asp |
| 725 | 730 | 735 |
| Leu Asn Ala Arg Val Lys | Ser Val Arg Glu Ala Ala | Glu Arg Met Ala |
| 740 | 745 | 750 |
| Phe Asn Met Pro Val Gln | Gly Thr Ala Ala Asp Leu | Met Lys Leu Ala |
| 755 | 760 | 765 |
| Met Val Lys Leu Phe Pro | Arg Leu Arg Glu Met Gly | Ala Arg Met Leu |
| 770 | 775 | 780 |
| Leu Gln Val Ala Asn Glu | Leu Leu Leu Glu Ala Pro | Gln Ala Arg Ala |
| 785 | 790 | 795 |
| Glu Glu Val Ala Ala Leu | Ala Lys Glu Ala Met Glu | Lys Ala Tyr Pro |
| 805 | 810 | 815 |
| Leu Ala Val Pro Leu Glu | Val Glu Val Gly Met Gly | Glu Asp Trp Leu |
| 820 | 825 | 830 |
| Ser Ala Lys Gly | | |
| 835 | | |

<210> 413
 <211> 42
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<220>
 <221> n
 <222> (19)..(21)
 <223> n is any base a,t,c, or g.

<400> 413
 caggagctta ggcttccnn nttgaagaag acgaagaaga ca

42

<210> 414
 <211> 18
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 414
cctaagctcg tcaaagag

18

<210> 415
<211> 2508
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 415
atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
gaaccggtgc aggcggtcta cggtctcgcc aagagcctcc tcaaggccct gaaggaggac 180
gggtacaagg ccgtcttcgt ggtctttgac gccaaaggccc cctccttcg ccacgaggcc 240
tacgaggcct acaaggcggg gagggccccc acccccagagg acttcccccg gcagctcgcc 300
ctcatcaagg agctggtgga cctcctgggg ttaccgcgc tcgagggtccc cggctacgag 360
gcggaacgacg ttctcgccac cctggccaag aaggcggaag aggaggggta cgagggtgcgc 420
atcctcaccg ccgaccgga cctctaccaa ctctctccg acccggtcgc cgtcctccac 480
cccaggggccc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag 540
cagtgggtgg acttccgcgc cctcgtgggg gaccctccg acaacctccc cgggggtcaag 600
ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc 660
ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg 720
gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctccc cctggagggtg 780
gacctcggcc aggggcggga gcccgaccgg gaggggctta gggccttcct ggagaggctg 840
gagttcggca gcctcctcca cgagttcggc ctctggagg ccccgcccc cctggaggag 900
gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctccc ccccgagccc 960
atgtgggcgg agcttaaagc cctggccgcc tgcaggggcg gccgctgca ccgggcagca 1020
gacccttggt cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc 1080
gccgtcttgg cctcgaggga ggggctagac ctctgcccc gggacgacct catgctcctc 1140
gcctacctcc tggaccttc gaacaccacc cccgaggggg tggcgcgggc ctacgggggg 1200
gagtggacgg aggacggcg ccaccgggccc ctctctcgg agaggctcca tcggaacctc 1260
cttaagcgcc tcgaggggga ggagaagctc ctttggtctt accacgaggt ggaaaagccc 1320

| | | | | | | |
|------------|------------|------------|-------------|------------|-------------|------|
| tacgaggcct | acaaggcggg | gagggccccc | acccccgagg | acttcccccg | gcagctcgcc | 300 |
| ctcatcaagg | agctggtgga | cctcctgggg | tttaccgcgc | tcgaggtccc | cggtacgag | 360 |
| gcggacgacg | ttctcgccac | cctggccaag | aaggcggaaa | aggaggggta | cgaggtgcgc | 420 |
| atcctcaccg | ccgaccgga | cctctacca | ctcgtctccg | accgcgtcgc | cgtoctccac | 480 |
| cccagggggc | acctcatcac | cccggagtgg | ctttgggaga | agtacggcct | caggccggag | 540 |
| cagtgggtgg | acttccgcgc | cctcgtgggg | gacccctccg | acaacctccc | cggggtcaag | 600 |
| ggcatcgggg | agaagaccgc | cctcaagctc | ctcaaggagt | ggggaagcct | ggaaaacctc | 660 |
| ctcaagaacc | tggaccgggt | aaagccagaa | aacgtccggg | agaagatcaa | ggcccacctg | 720 |
| gaagacctca | ggctctcctt | ggagctctcc | cgggtgcgca | ccgacctccc | cctggagggtg | 780 |
| gacctcgccc | aggggcggga | gcccgaaccg | gaggggctta | gggccttcct | ggagaggctg | 840 |
| gagttcggca | gcctcctcca | cgagttcggc | ctcctggagg | ccccgcgcc | cctggaggag | 900 |
| gccccctggc | ccccgccgga | aggggccttc | gtgggcttcg | tcctctcccg | ccccgagccc | 960 |
| atgtgggcgg | agcttaaagc | cctggccgcc | tgcaggggcg | gccgcgtgca | ccgggcagca | 1020 |
| gaccccttgg | cggggctaaa | ggacctcaag | gaggtccggg | gcctcctcgc | caaggacctc | 1080 |
| gccgtcttgg | cctcgaggga | ggggctagac | ctcgtgcccc | gggacgaccc | catgctcctc | 1140 |
| gcctacctcc | tggacccttc | gaacaccacc | cccagggggg | tggcgcggcg | ctacgggggg | 1200 |
| gagtggacgg | aggacgccgc | ccaccgggcc | ctcctctcgg | agaggctcca | tcggaacctc | 1260 |
| cttaagcgcc | tcgaggggga | ggagaagctc | ctttggctct | accacgaggt | ggaaaagccc | 1320 |
| ctctcccggg | tcctggccca | tatggaggcc | accgggggtac | ggcgggacgt | ggcctacctt | 1380 |
| caggcccttt | ccctggagct | tgcggaggag | atccgccgcc | tcgaggagga | ggtcttccgc | 1440 |
| ttggcggggc | acccttcaa | cctcaactcc | cgggaccagc | tggaaagggg | gctctttgac | 1500 |
| gagcttaggc | ttcccagttt | gaagaagacg | aagaagacag | gcaagcgctc | caccagcgcc | 1560 |
| gcggtgctgg | aggccctacg | ggaggccccc | cccatcgtag | agaagatcct | ccagcaccgg | 1620 |
| gagctcacca | agctcaagaa | cacctacgtg | gacccctccc | caagcctcgt | ccacccgagg | 1680 |
| acgggcccgc | tccacacccg | cttcaaccag | acggccacgg | ccacggggag | gcttagtagc | 1740 |
| tcggacccca | acctgcagaa | catccccgtc | cgcacccctt | tgggccagag | gatccgccgg | 1800 |
| gccttcgtgg | ccgaggcggg | ttgggcgttg | gtggccctgg | actatagcca | gatagagctc | 1860 |
| cgcgtcctcg | cccacctctc | cggggacgaa | aacctgatca | gggtcttcca | ggaggggaag | 1920 |
| gacatccaca | cccagaccgc | aagctggatg | ttcggcgctc | ccccggaggc | cgtggacccc | 1980 |

ctgatgcgcc gggcggccaa gacggtgaac ttcggcgctcc tctacggcat gtccgcccac 2040
 aggtctctccc aggagcttgc catcccctac gaggaggcgg tggcctttat agagcgctac 2100
 ttccaaagct tccccaaggt gcgggcctgg atagaaaaga ccctggagga ggggaggaag 2160
 cggggctacg tggaaccct cttcggaaga aggcgctacg tgcccacact caacgcccg 2220
 gtgaagagcg tcaggaggc cgcggagcgc atggccttca acatgcccg ccagggcacc 2280
 gccgcccacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg 2340
 gcccgcatgc tctccaggt cgccaacgag ctctcctgg agggccccca agcgcggggc 2400
 gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatccct cgcggtgccc 2460
 ctggaggtgg aggtggggat gggggaggac tggctttccg ccaagggt 2508

<210> 418
 <211> 836
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 418

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Asn | Ser | Glu | Ala | Met | Leu | Pro | Leu | Phe | Glu | Pro | Lys | Gly | Arg | Val |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Leu | Leu | Val | Asp | Gly | His | His | Leu | Ala | Tyr | Arg | Thr | Phe | Phe | Ala | Leu |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Lys | Gly | Leu | Thr | Thr | Ser | Arg | Gly | Glu | Pro | Val | Gln | Ala | Val | Tyr | Gly |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Phe | Ala | Lys | Ser | Leu | Leu | Lys | Ala | Leu | Lys | Glu | Asp | Gly | Tyr | Lys | Ala |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| Val | Phe | Val | Val | Phe | Asp | Ala | Lys | Ala | Pro | Ser | Phe | Arg | His | Glu | Ala |
| 65 | | | | 70 | | | | | 75 | | | | | 80 | |
| Tyr | Glu | Ala | Tyr | Lys | Ala | Gly | Arg | Ala | Pro | Thr | Pro | Glu | Asp | Phe | Pro |
| | | | | 85 | | | | | 90 | | | | | 95 | |
| Arg | Gln | Leu | Ala | Leu | Ile | Lys | Glu | Leu | Val | Asp | Leu | Leu | Gly | Phe | Thr |
| | | 100 | | | | | | 105 | | | | | 110 | | |
| Arg | Leu | Glu | Val | Pro | Gly | Tyr | Glu | Ala | Asp | Asp | Val | Leu | Ala | Thr | Leu |
| | | 115 | | | | | 120 | | | | | 125 | | | |
| Ala | Lys | Lys | Ala | Glu | Lys | Glu | Gly | Tyr | Glu | Val | Arg | Ile | Leu | Thr | Ala |
| | 130 | | | | | 135 | | | | | 140 | | | | |

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asp | Arg | Asp | Leu | Tyr | Gln | Leu | Val | Ser | Asp | Arg | Val | Ala | Val | Leu | His | 145 | 150 | 155 | 160 |
| Pro | Glu | Gly | His | Leu | Ile | Thr | Pro | Glu | Trp | Leu | Trp | Glu | Lys | Tyr | Gly | 165 | 170 | 175 | |
| Leu | Arg | Pro | Glu | Gln | Trp | Val | Asp | Phe | Arg | Ala | Leu | Val | Gly | Asp | Pro | 180 | 185 | 190 | |
| Ser | Asp | Asn | Leu | Pro | Gly | Val | Lys | Gly | Ile | Gly | Glu | Lys | Thr | Ala | Leu | 195 | 200 | 205 | |
| Lys | Leu | Leu | Lys | Glu | Trp | Gly | Ser | Leu | Glu | Asn | Leu | Leu | Lys | Asn | Leu | 210 | 215 | 220 | |
| Asp | Arg | Val | Lys | Pro | Glu | Asn | Val | Arg | Glu | Lys | Ile | Lys | Ala | His | Leu | 225 | 230 | 235 | 240 |
| Glu | Asp | Leu | Arg | Leu | Ser | Leu | Glu | Leu | Ser | Arg | Val | Arg | Thr | Asp | Leu | 245 | 250 | 255 | |
| Pro | Leu | Glu | Val | Asp | Leu | Ala | Gln | Gly | Arg | Glu | Pro | Asp | Arg | Glu | Gly | 260 | 265 | 270 | |
| Leu | Arg | Ala | Phe | Leu | Glu | Arg | Leu | Glu | Phe | Gly | Ser | Leu | Leu | His | Glu | 275 | 280 | 285 | |
| Phe | Gly | Leu | Leu | Glu | Ala | Pro | Ala | Pro | Leu | Glu | Glu | Ala | Pro | Trp | Pro | 290 | 295 | 300 | |
| Pro | Pro | Glu | Gly | Ala | Phe | Val | Gly | Phe | Val | Leu | Ser | Arg | Pro | Glu | Pro | 305 | 310 | 315 | 320 |
| Met | Trp | Ala | Glu | Leu | Lys | Ala | Leu | Ala | Ala | Cys | Arg | Gly | Gly | Arg | Val | 325 | 330 | 335 | |
| His | Arg | Ala | Ala | Asp | Pro | Leu | Ala | Gly | Leu | Lys | Asp | Leu | Lys | Glu | Val | 340 | 345 | 350 | |
| Arg | Gly | Leu | Leu | Ala | Lys | Asp | Leu | Ala | Val | Leu | Ala | Ser | Arg | Glu | Gly | 355 | 360 | 365 | |
| Leu | Asp | Leu | Val | Pro | Gly | Asp | Asp | Pro | Met | Leu | Leu | Ala | Tyr | Leu | Leu | 370 | 375 | 380 | |
| Asp | Pro | Ser | Asn | Thr | Thr | Pro | Glu | Gly | Val | Ala | Arg | Arg | Tyr | Gly | Gly | 385 | 390 | 395 | 400 |
| Glu | Trp | Thr | Glu | Asp | Ala | Ala | His | Arg | Ala | Leu | Leu | Ser | Glu | Arg | Leu | 405 | 410 | 415 | |
| His | Arg | Asn | Leu | Leu | Lys | Arg | Leu | Glu | Gly | Glu | Glu | Lys | Leu | Leu | Trp | 420 | 425 | 430 | |
| Leu | Tyr | His | Glu | Val | Glu | Lys | Pro | Leu | Ser | Arg | Val | Leu | Ala | His | Met | 435 | 440 | 445 | |

Glu Ala Thr Gly Val Arg Arg Asp Val Ala Tyr Leu Gln Ala Leu Ser
 450 455 460
 Leu Glu Leu Ala Glu Glu Ile Arg Arg Leu Glu Glu Glu Val Phe Arg
 465 470 475 480
 Leu Ala Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg
 485 490 495
 Val Leu Phe Asp Glu Leu Arg Leu Pro Ser Leu Lys Lys Thr Lys Lys
 500 505 510
 Thr Gly Lys Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu
 515 520 525
 Ala His Pro Ile Val Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys
 530 535 540
 Leu Lys Asn Thr Tyr Val Asp Pro Leu Pro Ser Leu Val His Pro Arg
 545 550 555 560
 Thr Gly Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly
 565 570 575
 Arg Leu Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr
 580 585 590
 Pro Leu Gly Gln Arg Ile Arg Arg Ala Phe Val Ala Glu Ala Gly Trp
 595 600 605
 Ala Leu Val Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala
 610 615 620
 His Leu Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys
 625 630 635 640
 Asp Ile His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu
 645 650 655
 Ala Val Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly
 660 665 670
 Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile
 675 680 685
 Pro Tyr Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe
 690 695 700
 Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys
 705 710 715 720
 Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp
 725 730 735
 Leu Asn Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala
 740 745 750

gccccctggc ccccgccgga aggggccttc gtgggcttcg tctctctccg ccccgagccc 960
atgtggggcg agcttaaagc cctggccgcc tgcagggcg gccgcgtgca ccgggcagca 1020
gaccccttgg cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc 1080
gcggtcttgg cctcgaggga ggggctagac ctctgtccccg gggacgaccc catgctcctc 1140
gcctacctcc tggaccttc gaacaccacc cccgaggggg tggcgcgcg ctacgggggg 1200
gagtggacgg aggacgcgc ccaccgggcc ctctctcgg agaggctcca tcggaacctc 1260
cttaagcgcc tcgaggggga ggagaagctc ctttggtct accacgaggt ggaaaagccc 1320
ctctccccgg tcttgccca tatggaggcc accggggtac ggcgggacgt ggcctacctt 1380
caggcccttt ccctggagct tgcggaggag atccgccgcc tcgaggagga ggtcttcgc 1440
ttggcgggcc accccttcaa cctcaactcc cgggaccagc tggaaagggg gctctttgac 1500
gagcttaggc ttccgcctt gaagaagacg aagaagacag gcaagcgctc caccggtgcc 1560
gcggtgctgg aggcctacg ggaggccac cccatcgtgg agaagatcct ccagcaccgg 1620
gagtcacca agctcaagaa cacctacgtg gacccctcc caagcctcgt ccacccgagg 1680
acgggcccgc tccacaccgc cttcaaccag acggccacgg ccacggggag gcttagtagc 1740
tccgaccca acctgcagaa catccccgtc cgcacccct tgggccagag gatccgcgg 1800
gccttcgtgg ccgaggcggg ttggcggtg gtggccctgg actatagcca gatagagctc 1860
cgcgtcctcg cccacctctc cggggacgaa aacctgatca gggctctcca ggaggggaag 1920
gacatccaca cccagaccgc aagctggatg ttcggcgtcc ccccgaggc cgtggacccc 1980
ctgatgcgcc gggcgcccaa gacgtgaac ttcggcgtcc tctacggcat gtccgcccac 2040
aggctctccc aggagcttgc catcccctac gaggaggcgg tggcctttat agagcgctac 2100
ttccaaagct tcccaagggt gcgggcctgg atagaaaaga ccctggagga ggggaggaag 2160
cggggctacg tggaaaccct cttcggaaga aggcgctacg tgcccgacct caacgcccgg 2220
gtgaagagcg tcaggaggc cgcggagcgc atggccttca acatgcccgt ccagggcacc 2280
gccgcgacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg 2340
gcccgcacgc tctccaggt cgccaacgag ctctccttgg agggccccc agcgcgggcc 2400
gaggagggtg cggttttggc caaggaggcc atggagaagg cctatcccct cgccgtgccc 2460
ctggagggtg aggtggggat gggggaggac tggctttccg ccaagggg 2508

<210> 420

<211> 836

<212> PRT
<213> Artificial

<220>
<223> Synthetic

<400> 420

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Asn | Ser | Glu | Ala | Met | Leu | Pro | Leu | Phe | Glu | Pro | Lys | Gly | Arg | Val |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Leu | Leu | Val | Asp | Gly | His | His | Leu | Ala | Tyr | Arg | Thr | Phe | Phe | Ala | Leu |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Lys | Gly | Leu | Thr | Thr | Ser | Arg | Gly | Glu | Pro | Val | Gln | Ala | Val | Tyr | Gly |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Phe | Ala | Lys | Ser | Leu | Leu | Lys | Ala | Leu | Lys | Glu | Asp | Gly | Tyr | Lys | Ala |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| Val | Phe | Val | Val | Phe | Asp | Ala | Lys | Ala | Pro | Ser | Phe | Arg | His | Glu | Ala |
| 65 | | | | | 70 | | | | | 75 | | | | 80 | |
| Tyr | Glu | Ala | Tyr | Lys | Ala | Gly | Arg | Ala | Pro | Thr | Pro | Glu | Asp | Phe | Pro |
| | | | | 85 | | | | | 90 | | | | | 95 | |
| Arg | Gln | Leu | Ala | Leu | Ile | Lys | Glu | Leu | Val | Asp | Leu | Leu | Gly | Phe | Thr |
| | | 100 | | | | | | 105 | | | | | 110 | | |
| Arg | Leu | Glu | Val | Pro | Gly | Tyr | Glu | Ala | Asp | Asp | Val | Leu | Ala | Thr | Leu |
| | 115 | | | | | | 120 | | | | | 125 | | | |
| Ala | Lys | Lys | Ala | Glu | Lys | Glu | Gly | Tyr | Glu | Val | Arg | Ile | Leu | Thr | Ala |
| | 130 | | | | | 135 | | | | | 140 | | | | |
| Asp | Arg | Asp | Leu | Tyr | Gln | Leu | Val | Ser | Asp | Arg | Val | Ala | Val | Leu | His |
| 145 | | | | | 150 | | | | 155 | | | | | 160 | |
| Pro | Glu | Gly | His | Leu | Ile | Thr | Pro | Glu | Trp | Leu | Trp | Glu | Lys | Tyr | Gly |
| | | | | 165 | | | | 170 | | | | | 175 | | |
| Leu | Arg | Pro | Glu | Gln | Trp | Val | Asp | Phe | Arg | Ala | Leu | Val | Gly | Asp | Pro |
| | | 180 | | | | | | 185 | | | | | 190 | | |
| Ser | Asp | Asn | Leu | Pro | Gly | Val | Lys | Gly | Ile | Gly | Glu | Lys | Thr | Ala | Leu |
| | 195 | | | | | | 200 | | | | | 205 | | | |
| Lys | Leu | Leu | Lys | Glu | Trp | Gly | Ser | Leu | Glu | Asn | Leu | Leu | Lys | Asn | Leu |
| | 210 | | | | | 215 | | | | | 220 | | | | |
| Asp | Arg | Val | Lys | Pro | Glu | Asn | Val | Arg | Glu | Lys | Ile | Lys | Ala | His | Leu |
| 225 | | | | | 230 | | | | | 235 | | | | 240 | |
| Glu | Asp | Leu | Arg | Leu | Ser | Leu | Glu | Leu | Ser | Arg | Val | Arg | Thr | Asp | Leu |
| | | | | 245 | | | | 250 | | | | | | 255 | |
| Pro | Leu | Glu | Val | Asp | Leu | Ala | Gln | Gly | Arg | Glu | Pro | Asp | Arg | Glu | Gly |

| 260 | 265 | 270 |
|-----------------------------|-------------------------|-----------------|
| Leu Arg Ala Phe Leu Glu Arg | Leu Glu Phe Gly Ser | Leu Leu His Glu |
| 275 | 280 | 285 |
| Phe Gly Leu Leu Glu Ala Pro | Ala Pro Leu Glu Glu Ala | Pro Trp Pro |
| 290 | 295 | 300 |
| Pro Pro Glu Gly Ala Phe Val | Gly Phe Val Leu Ser Arg | Pro Glu Pro |
| 305 | 310 | 315 |
| Met Trp Ala Glu Leu Lys Ala | Leu Ala Ala Cys Arg Gly | Gly Arg Val |
| 325 | 330 | 335 |
| His Arg Ala Ala Asp Pro Leu | Ala Gly Leu Lys Asp Leu | Lys Glu Val |
| 340 | 345 | 350 |
| Arg Gly Leu Leu Ala Lys Asp | Leu Ala Val Leu Ala Ser | Arg Glu Gly |
| 355 | 360 | 365 |
| Leu Asp Leu Val Pro Gly Asp | Asp Pro Met Leu Leu Ala | Tyr Leu Leu |
| 370 | 375 | 380 |
| Asp Pro Ser Asn Thr Thr Pro | Glu Gly Val Ala Arg Arg | Tyr Gly Gly |
| 385 | 390 | 395 |
| Glu Trp Thr Glu Asp Ala Ala | His Arg Ala Leu Leu Ser | Glu Arg Leu |
| 405 | 410 | 415 |
| His Arg Asn Leu Leu Lys Arg | Leu Glu Gly Glu Glu Lys | Leu Leu Trp |
| 420 | 425 | 430 |
| Leu Tyr His Glu Val Glu Lys | Pro Leu Ser Arg Val Leu | Ala His Met |
| 435 | 440 | 445 |
| Glu Ala Thr Gly Val Arg Arg | Asp Val Ala Tyr Leu Gln | Ala Leu Ser |
| 450 | 455 | 460 |
| Leu Glu Leu Ala Glu Glu Ile | Arg Arg Leu Glu Glu Glu | Val Phe Arg |
| 465 | 470 | 475 |
| Leu Ala Gly His Pro Phe Asn | Leu Asn Ser Arg Asp Gln | Leu Glu Arg |
| 485 | 490 | 495 |
| Val Leu Phe Asp Glu Leu Arg | Leu Pro Ala Leu Lys Lys | Thr Lys Lys |
| 500 | 505 | 510 |
| Thr Gly Lys Arg Ser Thr Gly | Ala Ala Val Leu Glu Ala | Leu Arg Glu |
| 515 | 520 | 525 |
| Ala His Pro Ile Val Glu Lys | Ile Leu Gln His Arg Glu | Leu Thr Lys |
| 530 | 535 | 540 |
| Leu Lys Asn Thr Tyr Val Asp | Pro Leu Pro Ser Leu Val | His Pro Arg |
| 545 | 550 | 555 |
| Thr Gly Arg Leu His Thr Arg | Phe Asn Gln Thr Ala Thr | Ala Thr Gly |

<220>
<223> Synthetic

<220>
<221> n
<222> (16)..(18)
<223> n is any base a,t,c, or g.

<400> 421
ggcaagcgct ccaccnnngc cgcggtgctg gaggcctac gg

42

<210> 422
<211> 15
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 422
ggtggagcgc ttgcc

15

<210> 423
<211> 2508
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 423
atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
ggccaccacc tggcctaccg cacctttcttc gccctgaagg gcctcaccac gagccggggc 120
gaaccggtgc aggcggtcta cggtctcgcc aagagcctcc tcaaggccct gaaggaggac 180
gggtacaagg ccgtcttcgt ggtctttgac gccaaaggccc cctccttcctg ccacgaggcc 240
tacgaggcct acaaggcggg gagggccccc acccccgagg acttcccccg gcagctcgcc 300
ctcatcaagg agctggtgga cctcctgggg tttaccgcgc tcgaggtccc cggctacgag 360
gcggacgacg ttctcgccac cctggccaag aaggcggaag aggaggggta cgaggtgcgc 420
atcctcaccg ccgaccgga cctctaccaa ctcgctctcc accgcgtcgc cgtcctccac 480
cccgagggcc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag 540
cagtgggtgg acttccgcgc cctcgtgggg gaccctcccg acaacctccc cgggggtcaag 600
ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc 660

| | | | | | | |
|------------|------------|------------|-------------|------------|-------------|------|
| ctcaagaacc | tggaccgggt | aaagccagaa | aacgtccggg | agaagatcaa | ggcccacctg | 720 |
| gaagacctca | ggctctcctt | ggagctctcc | cgggtgcgca | ccgacctccc | cctggagggtg | 780 |
| gacctcgccc | aggggcggga | gcccgaccgg | gaggggctta | gggccttcct | ggagaggctg | 840 |
| gagttcgga | gcctcctcca | cgagttcggc | ctcctggagg | cccccgcccc | cctggaggag | 900 |
| gccccctggc | ccccgcggga | aggggccttc | gtgggcttcg | tcctctcccg | ccccgagccc | 960 |
| atgtgggcg | agcttaaagc | cctggccgcc | tgcaggggcg | gccgcgtgca | ccgggcagca | 1020 |
| gacctcttgg | cggggctaaa | ggacctcaag | gaggtccggg | gcctcctcgc | caaggacctc | 1080 |
| gccgtcttgg | cctcgaggga | ggggctagac | ctcgtgcccc | gggacgaccc | catgctcctc | 1140 |
| gcctacctcc | tggaccttcc | gaacaccacc | cccagggggg | tggcgcgggc | ctacgggggg | 1200 |
| gagtggacgg | aggacgcgc | ccaccgggcc | ctcctctcgg | agaggctcca | tcggaacctc | 1260 |
| cttaagcgcc | tcgaggggga | ggagaagctc | ctttggctct | accacgaggt | ggaaaagccc | 1320 |
| ctctcccggg | tcctggccca | tatggaggcc | accgggggtac | ggcgggacgt | ggcctacctt | 1380 |
| caggcccttt | ccctggagct | tgcgaggagg | atccgccgcc | tcgaggagga | ggtcttccgc | 1440 |
| ttggcgggcc | acctcttcaa | cctcaactcc | cgggaccagc | tggaaagggg | gctctttgac | 1500 |
| gagcttaggc | ttcccgctt | gaagaagacg | aagaagacag | gcaagcgctc | caccagctta | 1560 |
| gcggtgctgg | aggccctacg | ggaggccac | cccatcgtag | agaagatcct | ccagcaccgg | 1620 |
| gagctcacca | agctcaagaa | cacctacgtg | gacccctccc | caagcctcgt | ccacccgagg | 1680 |
| acgggcccgc | tccacaccgc | cttcaaccag | acggccacgg | ccacggggag | gcttagtagc | 1740 |
| tccgacccca | acctgcagaa | catccccgtc | cgcacccctt | tgggccagag | gatccgcccg | 1800 |
| gccttcgtgg | ccgaggcggg | ttgggcgttg | gtggccctgg | actatagcca | gatagagctc | 1860 |
| cgcgtcctcg | cccacctctc | cggggacgaa | aacctgatca | gggtcttcca | ggaggggaag | 1920 |
| gacatccaca | cccagaccgc | aagctggatg | ttcggcgctc | ccccggaggc | cgtggacccc | 1980 |
| ctgatgcgcc | gggcgggcaa | gacggtgaac | ttcggcgctc | tctacggcat | gtccgcccac | 2040 |
| aggctctccc | aggagcttgc | catcccctac | gaggaggcgg | tggcctttat | agagcgctac | 2100 |
| ttccaaagct | tccccaaggt | gcgggcctgg | atagaaaaga | ccctggagga | ggggagggaag | 2160 |
| cggggctacg | tggaaaccct | cttcggaaga | aggcgctacg | tgcccgacct | caacgcccgg | 2220 |
| gtgaagagcg | tcaggagggc | cgcggagcgc | atggccttca | acatgcccgt | ccagggcacc | 2280 |
| gccgcgcacc | tcatgaagct | cgccatggtg | aagctcttcc | cccgcctccg | ggagatgggg | 2340 |
| gcccgcacgc | tcctccaggt | cgccaacgag | ctcctcctgg | aggcccccca | agcgcggggc | 2400 |

gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatcccct cgccgtgccc 2460

ctggaggtgg aggtggggat gggggaggac tggctttccg ccaaggggt 2508

<210> 424

<211> 836

<212> PRT

<213> Artificial

<220>

<223> Synthetic

<400> 424

Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
1 5 10 15

Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu
20 25 30

Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
35 40 45

Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala
50 55 60

Val Phe Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala
65 70 75 80

Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro
85 90 95

Arg Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr
100 105 110

Arg Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu
115 120 125

Ala Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala
130 135 140

Asp Arg Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His
145 150 155 160

Pro Glu Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly
165 170 175

Leu Arg Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro
180 185 190

Ser Asp Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu
195 200 205

Lys Leu Leu Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu
210 215 220

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asp | Arg | Val | Lys | Pro | Glu | Asn | Val | Arg | Glu | Lys | Ile | Lys | Ala | His | Leu | 225 | 230 | 235 | 240 |
| Glu | Asp | Leu | Arg | Leu | Ser | Leu | Glu | Leu | Ser | Arg | Val | Arg | Thr | Asp | Leu | 245 | 250 | 255 | |
| Pro | Leu | Glu | Val | Asp | Leu | Ala | Gln | Gly | Arg | Glu | Pro | Asp | Arg | Glu | Gly | 260 | 265 | 270 | |
| Leu | Arg | Ala | Phe | Leu | Glu | Arg | Leu | Glu | Phe | Gly | Ser | Leu | Leu | His | Glu | 275 | 280 | 285 | |
| Phe | Gly | Leu | Leu | Glu | Ala | Pro | Ala | Pro | Leu | Glu | Glu | Ala | Pro | Trp | Pro | 290 | 295 | 300 | |
| Pro | Pro | Glu | Gly | Ala | Phe | Val | Gly | Phe | Val | Leu | Ser | Arg | Pro | Glu | Pro | 305 | 310 | 315 | 320 |
| Met | Trp | Ala | Glu | Leu | Lys | Ala | Leu | Ala | Ala | Cys | Arg | Gly | Gly | Arg | Val | 325 | 330 | 335 | |
| His | Arg | Ala | Ala | Asp | Pro | Leu | Ala | Gly | Leu | Lys | Asp | Leu | Lys | Glu | Val | 340 | 345 | 350 | |
| Arg | Gly | Leu | Leu | Ala | Lys | Asp | Leu | Ala | Val | Leu | Ala | Ser | Arg | Glu | Gly | 355 | 360 | 365 | |
| Leu | Asp | Leu | Val | Pro | Gly | Asp | Asp | Pro | Met | Leu | Leu | Ala | Tyr | Leu | Leu | 370 | 375 | 380 | |
| Asp | Pro | Ser | Asn | Thr | Thr | Pro | Glu | Gly | Val | Ala | Arg | Arg | Tyr | Gly | Gly | 385 | 390 | 395 | 400 |
| Glu | Trp | Thr | Glu | Asp | Ala | Ala | His | Arg | Ala | Leu | Leu | Ser | Glu | Arg | Leu | 405 | 410 | 415 | |
| His | Arg | Asn | Leu | Leu | Lys | Arg | Leu | Glu | Gly | Glu | Glu | Lys | Leu | Leu | Trp | 420 | 425 | 430 | |
| Leu | Tyr | His | Glu | Val | Glu | Lys | Pro | Leu | Ser | Arg | Val | Leu | Ala | His | Met | 435 | 440 | 445 | |
| Glu | Ala | Thr | Gly | Val | Arg | Arg | Asp | Val | Ala | Tyr | Leu | Gln | Ala | Leu | Ser | 450 | 455 | 460 | |
| Leu | Glu | Leu | Ala | Glu | Glu | Ile | Arg | Arg | Leu | Glu | Glu | Glu | Val | Phe | Arg | 465 | 470 | 475 | 480 |
| Leu | Ala | Gly | His | Pro | Phe | Asn | Leu | Asn | Ser | Arg | Asp | Gln | Leu | Glu | Arg | 485 | 490 | 495 | |
| Val | Leu | Phe | Asp | Glu | Leu | Arg | Leu | Pro | Ala | Leu | Lys | Lys | Thr | Lys | Lys | 500 | 505 | 510 | |
| Thr | Gly | Lys | Arg | Ser | Thr | Ser | Leu | Ala | Val | Leu | Glu | Ala | Leu | Arg | Glu | 515 | 520 | 525 | |

| | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ala | His | Pro | Ile | Val | Glu | Lys | Ile | Leu | Gln | His | Arg | Glu | Leu | Thr | Lys | 530 | 535 | 540 |
| Leu | Lys | Asn | Thr | Tyr | Val | Asp | Pro | Leu | Pro | Ser | Leu | Val | His | Pro | Arg | 545 | 550 | 555 |
| Thr | Gly | Arg | Leu | His | Thr | Arg | Phe | Asn | Gln | Thr | Ala | Thr | Ala | Thr | Gly | 565 | 570 | 575 |
| Arg | Leu | Ser | Ser | Ser | Asp | Pro | Asn | Leu | Gln | Asn | Ile | Pro | Val | Arg | Thr | 580 | 585 | 590 |
| Pro | Leu | Gly | Gln | Arg | Ile | Arg | Arg | Ala | Phe | Val | Ala | Glu | Ala | Gly | Trp | 595 | 600 | 605 |
| Ala | Leu | Val | Ala | Leu | Asp | Tyr | Ser | Gln | Ile | Glu | Leu | Arg | Val | Leu | Ala | 610 | 615 | 620 |
| His | Leu | Ser | Gly | Asp | Glu | Asn | Leu | Ile | Arg | Val | Phe | Gln | Glu | Gly | Lys | 625 | 630 | 635 |
| Asp | Ile | His | Thr | Gln | Thr | Ala | Ser | Trp | Met | Phe | Gly | Val | Pro | Pro | Glu | 645 | 650 | 655 |
| Ala | Val | Asp | Pro | Leu | Met | Arg | Arg | Ala | Ala | Lys | Thr | Val | Asn | Phe | Gly | 660 | 665 | 670 |
| Val | Leu | Tyr | Gly | Met | Ser | Ala | His | Arg | Leu | Ser | Gln | Glu | Leu | Ala | Ile | 675 | 680 | 685 |
| Pro | Tyr | Glu | Glu | Ala | Val | Ala | Phe | Ile | Glu | Arg | Tyr | Phe | Gln | Ser | Phe | 690 | 695 | 700 |
| Pro | Lys | Val | Arg | Ala | Trp | Ile | Glu | Lys | Thr | Leu | Glu | Glu | Gly | Arg | Lys | 705 | 710 | 715 |
| Arg | Gly | Tyr | Val | Glu | Thr | Leu | Phe | Gly | Arg | Arg | Arg | Tyr | Val | Pro | Asp | 725 | 730 | 735 |
| Leu | Asn | Ala | Arg | Val | Lys | Ser | Val | Arg | Glu | Ala | Ala | Glu | Arg | Met | Ala | 740 | 745 | 750 |
| Phe | Asn | Met | Pro | Val | Gln | Gly | Thr | Ala | Ala | Asp | Leu | Met | Lys | Leu | Ala | 755 | 760 | 765 |
| Met | Val | Lys | Leu | Phe | Pro | Arg | Leu | Arg | Glu | Met | Gly | Ala | Arg | Met | Leu | 770 | 775 | 780 |
| Leu | Gln | Val | Ala | Asn | Glu | Leu | Leu | Leu | Glu | Ala | Pro | Gln | Ala | Arg | Ala | 785 | 790 | 795 |
| Glu | Glu | Val | Ala | Ala | Leu | Ala | Lys | Glu | Ala | Met | Glu | Lys | Ala | Tyr | Pro | 805 | 810 | 815 |
| Leu | Ala | Val | Pro | Leu | Glu | Val | Glu | Val | Gly | Met | Gly | Glu | Asp | Trp | Leu | 820 | 825 | 830 |

Ser Ala Lys Gly
835

<210> 425
<211> 42
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<220>
<221> n
<222> (19)..(21)
<223> n is any base a,t,c, or g.

<400> 425
ggcaagcgct ccaccagcnn ngcgggtgctg gaggcctac gg 42

<210> 426
<211> 2508
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 426
atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
gaaccgggtgc aggcgggteta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac 180
gggtacaagg ccgtcttcgt ggtctttgac gccaaaggccc cctccttcctg ccacgaggcc 240
tacgaggcct acaaggcggg gagggccccg acccccgagg acttcccccg gcagctcgcc 300
ctcatcaagg agctggtgga cctcctgggg tttaccgcc tgcaggtccc cggctacgag 360
gcggacgacg ttctcgccac cctggccaag aaggcggaaa aggaggggta cgaggtgcgc 420
atcctcaccg ccgaccgca cctctaccaa ctcgctctcc accgcgtcgc cgtcctccac 480
cccgagggcc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag 540
cagtgggtgg acttcgcgc cctcgtgggg gaccctccg acaacctccc cgggggtcaag 600
ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc 660
ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg 720
gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctccc cctggagggtg 780

| | |
|--|------|
| gacctcgccc aggggcggga gcccgaccgg gaggggctta gggccttcct ggagaggctg | 840 |
| gagttcggca gcctcctcca cgagttcggc ctcttgagg ccccgcccc cctggaggag | 900 |
| gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctcccg ccccgagccc | 960 |
| atgtgggcgg agcttaaagc cctggccgcc tgcaggggcg gccgcgtgca ccgggcagca | 1020 |
| gaccccttgg cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc | 1080 |
| gccgtcttgg cctcgaggga ggggctagac ctcgtgcccg gggacgacct catgctcctc | 1140 |
| gcctacctcc tggacccttc gaacaccacc cccgaggggg tggcgcgggc ctacgggggg | 1200 |
| gagtggacgg aggacgccgc ccaccgggcc ctctctcgg agaggctcca tcggaacctc | 1260 |
| cttaagcgcc tcgaggggga ggagaagctc ctttggtctc accacgaggt ggaaaagccc | 1320 |
| ctctcccggg tcctggccca tatggaggcc accgggggtac ggcgggacgt ggcctacctt | 1380 |
| ccctggagct tgcggaggag atccgccgc tcgaggagga ggtcttcgc | 1440 |
| ttggcgggcc accccttcaa cctcaactcc cgggaccagc tggaaagggg gctctttgac | 1500 |
| gagcttaggc ttccgcctt gaagaagacg aagaagacag gcaagcgctc caccagccgt | 1560 |
| gcggtgctgg aggccttacg ggaggccac cccatcgtgg agaagatcct ccagcaccgg | 1620 |
| gagctcacca agctcaagaa cacctacgtg gacccctcc caagcctcgt ccacccgagg | 1680 |
| acgggcgcgc tccacacccg cttcaaccag acggccacgg ccacggggag gcttagtagc | 1740 |
| tccgaccca acctgcagaa catccccgtc cgcacccct tgggcccagag gatccgcgg | 1800 |
| gccttcgtgg ccgaggcggg ttgggcgttg gtggcccttg actatagcca gatagagctc | 1860 |
| cgcgtcctcg cccacctctc cggggacgaa aacctgatca gggctctcca ggaggggaag | 1920 |
| gacatccaca cccagaccgc aagctggatg ttcggcgctc ccccgaggc cgtggacccc | 1980 |
| ctgatgcgc gggcgccaa gacggtgaac ttcggcgctc tctacggcat gtccgcccac | 2040 |
| aggctctccc aggagcttgc catcccctac gaggaggcgg tggcctttat agagcgctac | 2100 |
| ttcaaagct tccccagggt gcgggccttg atagaaaaga ccctggagga ggggaggaag | 2160 |
| cggggctacg tgaaaaccct cttcggaaga aggcgctacg tgcccgaact caacgcccgg | 2220 |
| gtgaagagcg tcagggaggc cgcggagcgc atggccttca acatgcccgt ccagggcacc | 2280 |
| gccgcccacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg | 2340 |
| gcccgcacgc tcctccaggt cgccaacgag ctctcctcgg agggccccca agcgcggggc | 2400 |
| gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatcccct cgccgtgccc | 2460 |
| ctggaggtgg aggtggggat gggggaggac tggctttccg ccaagggt | 2508 |

<210> 427
 <211> 836
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 427

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Asn | Ser | Glu | Ala | Met | Leu | Pro | Leu | Phe | Glu | Pro | Lys | Gly | Arg | Val | 1 | 5 | 10 | 15 |
| Leu | Leu | Val | Asp | Gly | His | His | Leu | Ala | Tyr | Arg | Thr | Phe | Phe | Ala | Leu | 20 | 25 | 30 | |
| Lys | Gly | Leu | Thr | Thr | Ser | Arg | Gly | Glu | Pro | Val | Gln | Ala | Val | Tyr | Gly | 35 | 40 | 45 | |
| Phe | Ala | Lys | Ser | Leu | Leu | Lys | Ala | Leu | Lys | Glu | Asp | Gly | Tyr | Lys | Ala | 50 | 55 | 60 | |
| Val | Phe | Val | Val | Phe | Asp | Ala | Lys | Ala | Pro | Ser | Phe | Arg | His | Glu | Ala | 65 | 70 | 75 | 80 |
| Tyr | Glu | Ala | Tyr | Lys | Ala | Gly | Arg | Ala | Pro | Thr | Pro | Glu | Asp | Phe | Pro | 85 | 90 | 95 | |
| Arg | Gln | Leu | Ala | Leu | Ile | Lys | Glu | Leu | Val | Asp | Leu | Leu | Gly | Phe | Thr | 100 | 105 | 110 | |
| Arg | Leu | Glu | Val | Pro | Gly | Tyr | Glu | Ala | Asp | Asp | Val | Leu | Ala | Thr | Leu | 115 | 120 | 125 | |
| Ala | Lys | Lys | Ala | Glu | Lys | Glu | Gly | Tyr | Glu | Val | Arg | Ile | Leu | Thr | Ala | 130 | 135 | 140 | |
| Asp | Arg | Asp | Leu | Tyr | Gln | Leu | Val | Ser | Asp | Arg | Val | Ala | Val | Leu | His | 145 | 150 | 155 | 160 |
| Pro | Glu | Gly | His | Leu | Ile | Thr | Pro | Glu | Trp | Leu | Trp | Glu | Lys | Tyr | Gly | 165 | 170 | 175 | |
| Leu | Arg | Pro | Glu | Gln | Trp | Val | Asp | Phe | Arg | Ala | Leu | Val | Gly | Asp | Pro | 180 | 185 | 190 | |
| Ser | Asp | Asn | Leu | Pro | Gly | Val | Lys | Gly | Ile | Gly | Glu | Lys | Thr | Ala | Leu | 195 | 200 | 205 | |
| Lys | Leu | Leu | Lys | Glu | Trp | Gly | Ser | Leu | Glu | Asn | Leu | Leu | Lys | Asn | Leu | 210 | 215 | 220 | |
| Asp | Arg | Val | Lys | Pro | Glu | Asn | Val | Arg | Glu | Lys | Ile | Lys | Ala | His | Leu | 225 | 230 | 235 | 240 |
| Glu | Asp | Leu | Arg | Leu | Ser | Leu | Glu | Leu | Ser | Arg | Val | Arg | Thr | Asp | Leu | | | | |

Thr Gly Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly
565 570 575

Arg Leu Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr
580 585 590

Pro Leu Gly Gln Arg Ile Arg Arg Ala Phe Val Ala Glu Ala Gly Trp
595 600 605

Ala Leu Val Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala
610 615 620

His Leu Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys
625 630 635 640

Asp Ile His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu
645 650 655

Ala Val Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly
660 665 670

Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile
675 680 685

Pro Tyr Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe
690 695 700

Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys
705 710 715 720

Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp
725 730 735

Leu Asn Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala
740 745 750

Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala
755 760 765

Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu
770 775 780

Leu Gln Val Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala
785 790 795 800

Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro
805 810 815

Leu Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu
820 825 830

Ser Ala Lys Gly
835

<210> 428
<211> 2499

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 428

| | |
|--|------|
| atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc | 60 |
| caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag | 120 |
| ccggtgcagg cggcttacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg | 180 |
| gacgcggtga tcgtggtctt tgacgccaaag gccccctcct tccgccacga ggcctacggg | 240 |
| gggtacaagg cgggccgggc cccacgccg gaggactttc cccggcaact cgccctcatc | 300 |
| aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac | 360 |
| gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc | 420 |
| accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccag | 480 |
| gggtacctca tcaccccggc ctggccttgg gaaaagtacg gcctgaggcc cgaccagtgg | 540 |
| gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc | 600 |
| ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag | 660 |
| aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg | 720 |
| aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggagggt ggacttcgcc | 780 |
| aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc | 840 |
| agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg | 900 |
| ccccgcggg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtggggc | 960 |
| gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat | 1020 |
| aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg | 1080 |
| gccctgaggg aaggccttgg cctcccggcc ggcgacgacc ccatgctcct cgcctacctc | 1140 |
| ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg | 1200 |
| gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttaagagg | 1260 |
| cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct | 1320 |
| gtcctggccc atatggaggc cacgggggtg cgccgggacg tggcctatct cagggccttg | 1380 |
| tccctggagg tggccgagga gatcgccgcg ctcgaggccg aggtcttccg cctggccggc | 1440 |
| caccccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg | 1500 |

```

cttcccccca tcaagaagac gcaaaaagacc ggcaagegct ccaccagcgc cgccgtcctg 1560
gagggccctcc gcgaggccca ccccatcggtg gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgaccccttg ccggacctca tccaccccag gacggggccgc 1680
ctccacacccc gcttcaacca gacggccaacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggcccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct caggggtgctg 1860
gcccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggagc ggacatccac 1920
acggagaccg ccagctggat gttcggcgtc cccggggagg ccgtggaccc cctgatgcgc 1980
cggggcgcca agaccatcaa cttcggggtc ctctacggca tgtcggccca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccaagg tgccggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
ctcatgaagc tggctatggt gaagctcttc ccaggctgg aggaaatggg ggccaggatg 2340
ctccttcagg tcgccaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
gcccggtgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggagggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggag 2499

```

```

<210> 429
<211> 833
<212> PRT
<213> Artificial

```

```

<220>
<223> Synthetic

```

```

<400> 429

```

```

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1          5          10          15

Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
          20          25          30

Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
          35          40          45

Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
          50          55          60

```


Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
675 680 685

Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
690 695 700

Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
705 710 715 720

Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
725 730 735

Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
740 745 750

Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
755 760 765

Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
770 775 780

Ala Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
785 790 795 800

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Glu

<210> 430
<211> 27
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 430
acgggggtgc gccgggacgt ggcctat

27

<210> 431
<211> 2508
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 431
atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
ggccaccacc tggcctaccg caccttcttc gcctgaagg gcctcaccac gagccggggg 120

| | | | | | | |
|------------|-------------|------------|-------------|------------|-------------|------|
| gaaccggtgc | aggcgggtcta | cggcttcgcc | aagagcctcc | tcaaggccct | gaaggaggac | 180 |
| gggtacaagg | cgtcttcgt | ggtctttgac | gccaaaggccc | cctccttcgc | ccacgaggcc | 240 |
| tacgaggcct | acaaggcggg | gagggccccc | acccccgagg | acttccccgc | gcagctcgcc | 300 |
| ctcatcaagg | agctggtgga | cctcctgggg | tttaccgcc | tcgaggctcc | cggctacgag | 360 |
| gcggacgacg | ttctcgccac | cctggccaag | aaggcggaaa | aggaggggta | cgagggtgcg | 420 |
| atcctcaccg | cgcaccgca | cctctaccaa | ctcgtctccg | accgcgtcgc | cgtcctccac | 480 |
| cccgagggcc | acctcatcac | cccggagtgg | ctttgggaga | agtacggcct | caggccggag | 540 |
| cagtgggtgg | acttcgcgc | cctcgtgggg | gacctctccg | acaacctccc | cgggggtcaag | 600 |
| ggcatcgggg | agaagaccgc | cctcaagctc | ctcaaggagt | ggggaagcct | ggaaaacctc | 660 |
| ctcaagaacc | tggaccgggt | aaagccagaa | aacgtccggg | agaagatcaa | ggccccactg | 720 |
| gaagacctca | ggctctcctt | ggagctctcc | cgggtgcgca | ccgacctccc | cctggagggtg | 780 |
| gacctcgccc | aggggcggga | gcccgaaccg | gaggggctta | gggccttcct | ggagaggctg | 840 |
| gagttcggca | gcctcctcca | cgagttcggc | ctcctggagg | ccccgcgcc | cctggaggag | 900 |
| gccccctggc | ccccgcggga | aggggccttc | gtgggcttcg | tcctctcccc | ccccgagccc | 960 |
| atgtgggcgg | agcttaaagc | cctggccgcc | tgcaggggcg | gccgcgtgca | ccgggcagca | 1020 |
| gacctcttgg | cggggctaaa | ggacctcaag | gaggtccggg | gcctcctcgc | caaggacctc | 1080 |
| gcgctcttgg | cctcgaggga | ggggctagac | ctcgtgcccc | gggacgacct | catgctcctc | 1140 |
| gcctacctcc | tggaccttc | gaacaccacc | ccccgagggg | tggcgcggcg | ctacgggggg | 1200 |
| gagtggaagg | aggacgccgc | ccaccgggcc | ctcctctcgg | agaggctcca | tcggaacctc | 1260 |
| cttaagcgcc | tcgaggggga | ggagaagctc | ctttggctct | accacgaggt | ggaaaagccc | 1320 |
| ctctccccgg | tcctggccca | tatggaggcc | accggggtag | ggcgggacgt | ggcctacctt | 1380 |
| caggcccttt | ccctggagct | tgcggaggag | atccgccgcc | tcgaggagga | ggtcttcgcg | 1440 |
| ttggcggggc | acctcttcaa | cctcaactcc | cgggaccagc | tggaaagggt | gctctttgac | 1500 |
| gagcttaggc | ttcccaagtt | gaagaagacg | aagaagacag | gcaagcgctc | caccagcgcc | 1560 |
| gcggtgctgg | aggccctacg | ggaggcccac | cccatcgtgg | agaagatcct | ccagcaccgg | 1620 |
| gagctcacca | agctcaagaa | cacctacgtg | gacccccctc | caagcctcgt | ccacccgagg | 1680 |
| acgggcgcgc | tccacaccgc | cttcaaccag | acggccacgg | ccacggggag | gcttagtagc | 1740 |
| tccgacccca | acctgcagaa | catccccgtc | cgcacccctc | tgggccagag | gatccgcggg | 1800 |

gccttcgtgg ccgaggcggg ttgggcgttg gtggccctgg actatagcca gatagagctc 1860
 cgcgctcctcg ccacacctc cggggacgaa aacctgatca gggctttcca ggaggggaag 1920
 gacatccaca ccagaccgc aagctggatg ttcggcgctcc ccccgagggc cgtggacccc 1980
 ctgatgcgcc gggcgccaa gacgggtgaac ttcggcgctcc tctacggcat gtccgccccat 2040
 aggctctccc aggagcttgc catcccctac gaggaggcgg tggcctttat agagcgctac 2100
 ttccaaagct tcccaaaggt gggggcctgg atagaaaaga ccctggagga ggggaggaag 2160
 cggggctacg tggaaaccct cttcggaaga aggcgctacg tgcccacact caacgcccg 2220
 gtgaagagcg tcaggagggc cgcggagcgc atggccttca acatgcccg ccagggcacc 2280
 gccgccgacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg 2340
 gcccgcatgc tctccaggt cgccaacgag ctctccttgaggggccca agcgcgggcc 2400
 gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatcccct cgccgtgccc 2460
 ctggaggtgg aggtggggat gggggaggac tggctttccg ccaaggg 2508

<210> 432
 <211> 836
 <212> PRT
 <213> Artificial
 <220>
 <223> Synthetic
 <400> 432

Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
 1 5 10 15
 Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu
 20 25 30
 Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
 35 40 45
 Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala
 50 55 60
 Val Phe Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala
 65 70 75 80
 Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro
 85 90 95
 Arg Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr
 100 105 110
 Arg Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu

| | | |
|---|-----|-----|
| 115 | 120 | 125 |
| Ala Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala | | |
| 130 | 135 | 140 |
| Asp Arg Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His | | |
| 145 | 150 | 155 |
| Pro Glu Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly | | |
| | 165 | 170 |
| | | 175 |
| Leu Arg Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro | | |
| | 180 | 185 |
| | | 190 |
| Ser Asp Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu | | |
| | 195 | 200 |
| | | 205 |
| Lys Leu Leu Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu | | |
| | 210 | 215 |
| | | 220 |
| Asp Arg Val Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu | | |
| | 225 | 230 |
| | | 235 |
| Glu Asp Leu Arg Leu Ser Leu Glu Leu Ser Arg Val Arg Thr Asp Leu | | |
| | 245 | 250 |
| | | 255 |
| Pro Leu Glu Val Asp Leu Ala Gln Gly Arg Glu Pro Asp Arg Glu Gly | | |
| | 260 | 265 |
| | | 270 |
| Leu Arg Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu | | |
| | 275 | 280 |
| | | 285 |
| Phe Gly Leu Leu Glu Ala Pro Ala Pro Leu Glu Glu Ala Pro Trp Pro | | |
| | 290 | 295 |
| | | 300 |
| Pro Pro Glu Gly Ala Phe Val Gly Phe Val Leu Ser Arg Pro Glu Pro | | |
| | 305 | 310 |
| | | 315 |
| Met Trp Ala Glu Leu Lys Ala Leu Ala Ala Cys Arg Gly Gly Arg Val | | |
| | 325 | 330 |
| | | 335 |
| His Arg Ala Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val | | |
| | 340 | 345 |
| | | 350 |
| Arg Gly Leu Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly | | |
| | 355 | 360 |
| | | 365 |
| Leu Asp Leu Val Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu | | |
| | 370 | 375 |
| | | 380 |
| Asp Pro Ser Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly | | |
| | 385 | 390 |
| | | 395 |
| | | 400 |
| Glu Trp Thr Glu Asp Ala Ala His Arg Ala Leu Leu Ser Glu Arg Leu | | |
| | 405 | 410 |
| | | 415 |
| His Arg Asn Leu Leu Lys Arg Leu Glu Gly Glu Glu Lys Leu Leu Trp | | |

| | | |
|---|-----|-----|
| 725 | 730 | 735 |
| Leu Asn Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala | | |
| 740 | 745 | 750 |
| Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala | | |
| 755 | 760 | 765 |
| Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu | | |
| 770 | 775 | 780 |
| Leu Gln Val Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala | | |
| 785 | 790 | 795 |
| Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro | | |
| 805 | 810 | 815 |
| Leu Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu | | |
| 820 | 825 | 830 |
| Ser Ala Lys Gly | | |
| 835 | | |
| <210> 433 | | |
| <211> 36 | | |
| <212> DNA | | |
| <213> Artificial | | |
| <220> | | |
| <223> Synthetic | | |
| <400> 433 | | |
| cttaggcttc ccaagttgaa gaagacgaag aagaca | | 36 |
| <210> 434 | | |
| <211> 36 | | |
| <212> DNA | | |
| <213> Artificial | | |
| <220> | | |
| <223> Synthetic | | |
| <400> 434 | | |
| tgtcttcttc gtcttcttca acttggaag cctaag | | 36 |
| <210> 435 | | |
| <211> 2508 | | |
| <212> DNA | | |
| <213> Artificial | | |
| <220> | | |
| <223> Synthetic | | |
| <400> 435 | | |
| atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac | | 60 |

ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
gaaccggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac 180
gggtacaagg ccgtcttcgt ggtctttgac gccaaaggccc cctccttcg ccacgaggcc 240
tacgaggcct acaaggcggg gagggccccg acccccagagg acttcccccg gcagctcgcc 300
ctcatcaagg agctggtgga cctcctgggg ttaccgcgcc tcgaggctcc cggctacgag 360
gcggaacgacg ttctcgccac cctggccaag aaggcggaaa aggaggggta cgagggtgcgc 420
atcctcaccg ccgaccgcga cctctaccaa ctctgtctcc accgcgtcgc cgtcctccac 480
cccaggggcc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag 540
cagtgggtgg acttcgcgc cctcgtgggg gacccctccg acaacctccc cgggggtcaag 600
ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc 660
ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg 720
gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctccc cctggagggtg 780
gacctcgccc aggggcggga gcccgaccgg gaggggctta gggccttcct ggagaggctg 840
gagttcggca gcctcctcca cgagttcggc ctcttgagg ccccgcccc cctggaggag 900
gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctcccg ccccgagccc 960
atgtgggcgg agcttaaagc cctggccgcc tgcaggggcg gccgcgtgca ccgggcagca 1020
gaccccttgg cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc 1080
gccgtcttgg cctcgaggga ggggctagac ctctgcccc gggacgacct catgctcctc 1140
gcctacctcc tggacccttc gaacaccacc cccgaggggg tggcgcgggc ctacgggggg 1200
gagtggaacg aggacgccgc ccaccgggcc ctctctcgg agaggctcca tcggaacctc 1260
cttaagcgcc tcgaggggga ggagaagctc ctttggtct accacgagggt ggaaaagccc 1320
ctctcccggg tcctggccca tatggaggcc accggggtag ggcgggacgt ggcctacctt 1380
caggcccttt ccctggagct tgcggaggag atccgccgcc tcgaggagga ggtcttcgc 1440
ttggcggggc accccttcaa cctcaactcc cgggaccage tggaaagggt gctctttgac 1500
gagcttaggc ttccgcctt gaagaagacg aagaagacag gcaagcgctc caccagcgcc 1560
gcggtgctgg aggccttacg ggaggccac cccatcgtgg agaagatcct ccagcaccgg 1620
gagctacca agctcaagaa cacctacgtg gacccctcc caagcctcgt ccaccgagg 1680
acgggccgcc tccacaccg cttcaaccag acggccacgg ccacggggag gcttagtagc 1740

tccgacccca acctgcagaa catccccgtc cgcaccccct tgggccagag gatccgccgg 1800
 gccttcgtgg ccgaggcggg ttggggcgttg gtggcccttg actatagcca gatagagctc 1860
 cgcgtcctcg cccacctctc cggggacgaa aacctgatca gggctctcca ggaggggaag 1920
 gacatcgcca cccagaccgc aagctggatg ttcggcgctc ccccgagggc cgtggacccc 1980
 ctgatgcgcc gggcggccaa gacggtgaac ttcggcgctc tctacggcat gtccgcccat 2040
 aggctctccc aggagcttgc catcccctac gaggaggcgg tggcctttat agagcgctac 2100
 ttccaaagct tccccaaagt gcgggccttg atagaaaaga ccctggagga ggggaggaag 2160
 cggggctacg tggaaaccct cttcggaaga aggcgctacg tgcccgaact caacgcccgg 2220
 gtgaagagcg tcagggaggc cgcggagcgc atggccttca acatgcccggt ccagggcacc 2280
 gccgccgacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg 2340
 gcccgcatgc tcctccaggt cgccaacgag ctctcctcgg agggccccca agcgcggggc 2400
 gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatcccct cgccgtgccc 2460
 ctggaggtgg aggtggggat gggggaggac tggctttccg ccaagggt 2508

<210> 436
 <211> 836
 <212> PRT
 <213> Artificial
 <220>
 <223> Synthetic
 <400> 436

Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
 1 5 10 15
 Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu
 20 25 30
 Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
 35 40 45
 Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala
 50 55 60
 Val Phe Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala
 65 70 75 80
 Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro
 85 90 95
 Arg Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr
 100 105 110

Arg Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu
 115 120 125
 Ala Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala
 130 135 140
 Asp Arg Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His
 145 150 155 160
 Pro Glu Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly
 165 170 175
 Leu Arg Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro
 180 185 190
 Ser Asp Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu
 195 200 205
 Lys Leu Leu Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu
 210 215 220
 Asp Arg Val Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu
 225 230 235 240
 Glu Asp Leu Arg Leu Ser Leu Glu Leu Ser Arg Val Arg Thr Asp Leu
 245 250 255
 Pro Leu Glu Val Asp Leu Ala Gln Gly Arg Glu Pro Asp Arg Glu Gly
 260 265 270
 Leu Arg Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu
 275 280 285
 Phe Gly Leu Leu Glu Ala Pro Ala Pro Leu Glu Glu Ala Pro Trp Pro
 290 295 300
 Pro Pro Glu Gly Ala Phe Val Gly Phe Val Leu Ser Arg Pro Glu Pro
 305 310 315 320
 Met Trp Ala Glu Leu Lys Ala Leu Ala Ala Cys Arg Gly Gly Arg Val
 325 330 335
 His Arg Ala Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val
 340 345 350
 Arg Gly Leu Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly
 355 360 365
 Leu Asp Leu Val Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu
 370 375 380
 Asp Pro Ser Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly
 385 390 395 400
 Glu Trp Thr Glu Asp Ala Ala His Arg Ala Leu Leu Ser Glu Arg Leu
 405 410 415

Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp
725 730 735

Leu Asn Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala
740 745 750

Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala
755 760 765

Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu
770 775 780

Leu Gln Val Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala
785 790 795 800

Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro
805 810 815

Leu Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu
820 825 830

Ser Ala Lys Gly
835

<210> 437
<211> 33
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 437
gaggggaagg acatcgccac ccagaccgca agc

33

<210> 438
<211> 33
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 438
gcttgcggtc tgggtggcga tgccttccc etc

33

<210> 439
<211> 2508
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 439
atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
gaaccggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac 180
gggtacaagg cegtcttcgt ggtctttgac gccaaaggccc cctccttcg ccacgaggcc 240
tacgaggcct acaaggcggg gagggcccg acccccgagg acttcccccg gcagctcgcc 300
ctcatcaagg agctggtgga cctcctgggg ttaccgcgc tcgaggtccc cggctacgag 360
gcggacgacg ttctcgccac cctggccaag aaggcggaaa aggaggggta cgaggtgcgc 420
atcctcaccg ccgaccgga cctctaccaa ctcgctcgc accgcgtcgc cgtcctccac 480
cccgagggcc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag 540
cagtgggtgg acttcgcgc cctcgtgggg gacccctccg acaacctccc cggggtcaag 600
ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc 660
ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggccacctg 720
gaagacctca ggctctcctt ggagctctcc cgggtgcga ccgacctccc cctggaggtg 780
gacctcgccc aggggcggga gcccgaccgg gaggggctta gggccttcct ggagaggctg 840
gagttcgga gcctcctcca cgagttcggc ctctggagg ccccgcccc cctggaggag 900
gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctcccc ccccgagccc 960
atgtgggcgg agcttaaagc cctggccgcc tgcagggcg gccgcgtgca ccgggcagca 1020
gaccccttgg cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc 1080
gcgctcttgg cctcgaggga ggggctagac ctcggtcccc gggacgacct catgctcctc 1140
gcctacctcc tggaccttc gaacaccacc cccgaggggg tggcgcgcg ctacgggggg 1200
gagtggacgg aggacgccgc ccaccgggcc ctctctcgg agaggctcca tcggaacctc 1260
cttaagcgcc tcgaggggga ggagaagctc ctttggctct accacgaggt ggaaaagccc 1320
ctctcccggg tcctggccca tatggaggcc accggggtac ggcgggacgt ggctacctt 1380
caggcccttt ccctggagct tgcggaggag atccgccgcc tcgaggagga ggtcttcgc 1440
ttggcgggcc accccttcaa cctcaactcc cgggaccagc tggaaagggt gctctttgac 1500
gagcttaggc ttcccgctt gaagaagccg aagaagacag gcaagcgctc caccagcgcc 1560
gcggtgctgg aggcctacg ggaggccac cccatcgtgg agaagatcct ccagcaccgg 1620
gagctacca agctcaagaa cacctacgtg gacccctcc caagcctcgt ccacccgagg 1680

acggggccgcc tccacacccg cttcaaccag acggccacgg ccacggggag gcttagtagc 1740
 tccgacccca acctgcagaa catccccgtc cgcaccccct tgggccagag gatccgccgg 1800
 gccttcgtgg ccgaggcggg ttgggcgttg gtggccctgg actatagcca gatagagctc 1860
 cgcgctcctcg cccacctctc cggggacgaa aacctgatca gggctctcca ggaggggaag 1920
 gacatccaca cccagaccgc aagctggatg ttcggcgctc ccccgagggc cgtggacccc 1980
 ctgatgcgcc gggcggccaa gacggtgaac ttcggcgctc tctacggcat gtccgcccac 2040
 aggctctccc aggagcttgc catcccctac gaggaggcgg tggcctttat agagcgctac 2100
 ttccaaagct tcccaaggt gcgggcctgg atagaaaaga ccctggagga ggggaggaag 2160
 cggggctacg tggaaaccct cttcggaaga aggcgctacg tgcccgacct caacgcccgg 2220
 gtgaagagcg tcagggaggc cgcggagcgc atggccttca acatgcccgt ccagggcacc 2280
 gccgccgacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg 2340
 gcccgcatgc tcctccaggt cgccaacgag ctctctctgg agggccccca agcgcgggcc 2400
 gaggaggtgg cggttttggc caaggaggcc atggagaagg cctatcccct cgccgtgccc 2460
 ctggaggtgg aggtggggat gggggaggac tggctttccg ccaagggt 2508

<210> 440
 <211> 836
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 440

Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
1 5 10 15

Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu
20 25 30

Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
35 40 45

Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala
50 55 60

Val Phe Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala
65 70 75 80

Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro
85 90 95

Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys
705 710 715 720

Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp
725 730 735

Leu Asn Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala
740 745 750

Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala
755 760 765

Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu
770 775 780

Leu Gln Val Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala
785 790 795 800

Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro
805 810 815

Leu Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu
820 825 830

Ser Ala Lys Gly
835

<210> 441
<211> 33
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 441
cccgcccttga agaagccgaa gaagacaggc aag

33

<210> 442
<211> 33
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 442
cttgcctgtc ttcttcggct tcttcaaggc ggg

33

<210> 443
<211> 2499
<212> DNA
<213> Artificial

<220>

<223> Synthetic

<400> 443

| | |
|--|------|
| atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc | 60 |
| caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag | 120 |
| ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg | 180 |
| gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg | 240 |
| gggtacaagg cgggcccggc cccacgccg gaggactttc cccggcaact cgccctcatc | 300 |
| aaggagctgg tggacctcct ggggttcacg cgctctgagg tcccgggcta cgaggcggac | 360 |
| gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc | 420 |
| accgcccagaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccag | 480 |
| gggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg | 540 |
| gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc | 600 |
| ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag | 660 |
| aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg | 720 |
| aagctctcct gggacctggc caaggtgctc accgacctgc ccctggaggt ggacttcgcc | 780 |
| aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc | 840 |
| agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg | 900 |
| ccccgcggg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtggggc | 960 |
| gatcttctgg ccctggccgc cgccaggggc ggccgcgttt accgggcgga ggatcccttg | 1020 |
| gaggccttgc gggggccttg ggaggtgagg gggcttttgg ccaaggacct ggcggtgctg | 1080 |
| gccctgaggg aagggattgc cctggcaccg ggcgacgacc ccatgctcct cgcctacctc | 1140 |
| ctggatcctt ccaacaccgc cccgaagggt gtagcccgcc gctacggggg ggagtggacc | 1200 |
| gaggaggcgg gggaaagggc gctgctttcc gaaaggcttt acgccgccct cctgaagcgg | 1260 |
| cttaaggggg aggagaggct tctttggctt tacgaggagg tggaaaagcc cttttcgcg | 1320 |
| gtcctggccc acatggaggc cacgggggta cggttggatg tggcctactt aaaggccctt | 1380 |
| tccctggagg tggaggcggg gataaggcgc ttcgaggagg aggtccaccg cctggccggg | 1440 |
| catcctttca acctgaactc ccgggaccag ctggaaaggg tcctctttga cgagcttggg | 1500 |
| cttcccgcga tcaagaagac gaggaagacg ggcaagcgct ccaccagcgc cgccgttttg | 1560 |
| gaggccttgc gggaggctca tcccatcgtg gaccgcctcc ttcagtaccg ggagctttcc | 1620 |

| | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|--|
| 85 | | | | | | | | 90 | | | | | 95 | | | | |
| Leu | Ala | Leu | Ile | Lys | Glu | Leu | Val | Asp | Leu | Leu | Gly | Phe | Thr | Arg | Leu | | |
| | | | 100 | | | | | 105 | | | | | 110 | | | | |
| Glu | Val | Pro | Gly | Tyr | Glu | Ala | Asp | Asp | Val | Leu | Ala | Ser | Leu | Ala | Lys | | |
| | | 115 | | | | | 120 | | | | | 125 | | | | | |
| Lys | Ala | Glu | Lys | Glu | Gly | Tyr | Glu | Val | Arg | Ile | Leu | Thr | Ala | Asp | Lys | | |
| | 130 | | | | | 135 | | | | | 140 | | | | | | |
| Asp | Leu | Tyr | Gln | Leu | Leu | Ser | Asp | Arg | Ile | His | Val | Leu | His | Pro | Glu | | |
| 145 | | | | 150 | | | | | 155 | | | | | | 160 | | |
| Gly | Tyr | Leu | Ile | Thr | Pro | Ala | Trp | Leu | Trp | Glu | Lys | Tyr | Gly | Leu | Arg | | |
| | | | 165 | | | | | 170 | | | | | | 175 | | | |
| Pro | Asp | Gln | Trp | Ala | Asp | Tyr | Arg | Ala | Leu | Thr | Gly | Asp | Glu | Ser | Asp | | |
| | | 180 | | | | | 185 | | | | | | 190 | | | | |
| Asn | Leu | Pro | Gly | Val | Lys | Gly | Ile | Gly | Glu | Lys | Thr | Ala | Arg | Lys | Leu | | |
| | 195 | | | | | 200 | | | | | | 205 | | | | | |
| Leu | Glu | Glu | Trp | Gly | Ser | Leu | Glu | Ala | Leu | Leu | Lys | Asn | Leu | Asp | Arg | | |
| | 210 | | | | | 215 | | | | | 220 | | | | | | |
| Leu | Lys | Pro | Ala | Ile | Arg | Glu | Lys | Ile | Leu | Ala | His | Met | Asp | Asp | Leu | | |
| 225 | | | | 230 | | | | | | 235 | | | | | 240 | | |
| Lys | Leu | Ser | Trp | Asp | Leu | Ala | Lys | Val | Arg | Thr | Asp | Leu | Pro | Leu | Glu | | |
| | | | 245 | | | | | 250 | | | | | | 255 | | | |
| Val | Asp | Phe | Ala | Lys | Arg | Arg | Glu | Pro | Asp | Arg | Glu | Arg | Leu | Arg | Ala | | |
| | | 260 | | | | | 265 | | | | | | 270 | | | | |
| Phe | Leu | Glu | Arg | Leu | Glu | Phe | Gly | Ser | Leu | Leu | His | Glu | Phe | Gly | Leu | | |
| | 275 | | | | | 280 | | | | | 285 | | | | | | |
| Leu | Glu | Ser | Pro | Lys | Ala | Leu | Glu | Glu | Ala | Pro | Trp | Pro | Pro | Pro | Glu | | |
| | 290 | | | | | 295 | | | | | 300 | | | | | | |
| Gly | Ala | Phe | Val | Gly | Phe | Val | Leu | Ser | Arg | Lys | Glu | Pro | Met | Trp | Ala | | |
| 305 | | | | 310 | | | | | | 315 | | | | | 320 | | |
| Asp | Leu | Leu | Ala | Leu | Ala | Ala | Ala | Arg | Gly | Gly | Arg | Val | Tyr | Arg | Ala | | |
| | | | 325 | | | | | 330 | | | | | | 335 | | | |
| Glu | Asp | Pro | Leu | Glu | Ala | Leu | Arg | Gly | Leu | Gly | Glu | Val | Arg | Gly | Leu | | |
| | | 340 | | | | | 345 | | | | | | 350 | | | | |
| Leu | Ala | Lys | Asp | Leu | Ala | Val | Leu | Ala | Leu | Arg | Glu | Gly | Ile | Ala | Leu | | |
| | | 355 | | | | 360 | | | | | 365 | | | | | | |
| Ala | Pro | Gly | Asp | Asp | Pro | Met | Leu | Leu | Ala | Tyr | Leu | Leu | Asp | Pro | Ser | | |
| | 370 | | | | | 375 | | | | | 380 | | | | | | |
| Asn | Thr | Ala | Pro | Glu | Gly | Val | Ala | Arg | Arg | Tyr | Gly | Gly | Glu | Trp | Thr | | |

| | | | | | | |
|-----------------|-------------------------|-----------------------------|-----|-----|--|-----|
| 385 | | 390 | | 395 | | 400 |
| Glu Glu Ala Gly | Glu Arg Ala Leu Leu Ser | Glu Arg Leu Tyr Ala Ala | | | | |
| | 405 | | 410 | | | 415 |
| Leu Leu Lys Arg | Leu Lys Gly Glu Glu Arg | Leu Leu Trp Leu Tyr Glu | | | | |
| | 420 | | 425 | | | 430 |
| Glu Val Glu Lys | Pro Leu Ser Arg Val | Leu Ala His Met Glu Ala Thr | | | | |
| | 435 | | 440 | | | 445 |
| Gly Val Arg Leu | Asp Val Ala Tyr Leu Lys | Ala Leu Ser Leu Glu Val | | | | |
| | 450 | | 455 | | | 460 |
| Glu Ala Glu Ile | Arg Arg Phe Glu Glu Glu | Val His Arg Leu Ala Gly | | | | |
| | 465 | | 470 | | | 475 |
| His Pro Phe Asn | Leu Asn Ser Arg Asp | Gln Leu Glu Arg Val Ile Phe | | | | |
| | 485 | | 490 | | | 495 |
| Asp Glu Leu Gly | Leu Pro Ala Ile Lys Lys | Thr Arg Lys Thr Gly Lys | | | | |
| | 500 | | 505 | | | 510 |
| Arg Ser Thr Ser | Ala Ala Val Leu Glu Ala | Leu Arg Glu Ala His Pro | | | | |
| | 515 | | 520 | | | 525 |
| Ile Val Asp Arg | Ile Leu Gln Tyr Arg | Glu Leu Ser Lys Leu Lys Gly | | | | |
| | 530 | | 535 | | | 540 |
| Thr Tyr Ile Asp | Pro Leu Pro Ala Leu Val | His Pro Lys Thr Asn Arg | | | | |
| | 545 | | 550 | | | 555 |
| Leu His Thr Arg | Phe Asn Gln Thr Ala | Thr Ala Thr Gly Arg Leu Ser | | | | |
| | 565 | | 570 | | | 575 |
| Ser Ser Asp Pro | Asn Leu Gln Asn Ile | Pro Val Arg Thr Pro Leu Gly | | | | |
| | 580 | | 585 | | | 590 |
| Gln Arg Ile Arg | Arg Ala Phe Val Ala | Glu Glu Gly Trp Arg Leu Val | | | | |
| | 595 | | 600 | | | 605 |
| Val Leu Asp Tyr | Ser Gln Ile Glu Leu Arg | Val Leu Ala His Leu Ser | | | | |
| | 610 | | 615 | | | 620 |
| Gly Asp Glu Asn | Leu Ile Arg Val Phe | Gln Glu Gly Gln Asp Ile His | | | | |
| | 625 | | 630 | | | 635 |
| Thr Gln Thr Ala | Ser Trp Met Phe Gly | Val Pro Pro Glu Ala Val Asp | | | | |
| | 645 | | 650 | | | 655 |
| Ser Leu Met Arg | Arg Ala Ala Lys Thr | Ile Asn Phe Gly Val Leu Tyr | | | | |
| | 660 | | 665 | | | 670 |
| Gly Met Ser Ala | His Arg Leu Ser Gly | Glu Leu Ala Ile Pro Tyr Glu | | | | |
| | 675 | | 680 | | | 685 |
| Glu Ala Val Ala | Phe Ile Glu Arg Tyr | Phe Gln Ser Tyr Pro Lys Val | | | | |

| | | |
|---|-----|-------------|
| 690 | 695 | 700 |
| Arg Ala Trp Ile Glu Lys Thr Leu Ala Glu Gly Arg Glu Arg Gly Tyr | | |
| 705 | 710 | 715 720 |
| Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Ala Ser | | |
| | 725 | 730 735 |
| Arg Val Lys Ser Ile Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met | | |
| | 740 | 745 750 |
| Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys | | |
| | 755 | 760 765 |
| Leu Phe Pro Arg Leu Gln Glu Leu Gly Ala Arg Met Leu Leu Gln Val | | |
| | 770 | 775 780 |
| His Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Gln Ala Glu Glu Val | | |
| | 785 | 790 795 800 |
| Ala Gln Glu Ala Lys Arg Thr Met Glu Glu Val Trp Pro Leu Lys Val | | |
| | 805 | 810 815 |
| Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys | | |
| | 820 | 825 830 |

Ala

<210> 445
 <211> 2496
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

| | |
|--|-----|
| <400> 445 | |
| atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc | 60 |
| caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag | 120 |
| ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggcctcaa ggaggacggg | 180 |
| gacgcggtga tcgtggtctt tgacgccaa gccccctcct tccgccacga ggcctacggg | 240 |
| gggtacaagg cgggcccgggc cccacgccg gaggactttc cccggcaact cgccctcatc | 300 |
| aaggagctgg tggacctcct ggggttcacg cgcctcgagg tcccgggcta cgaggcggac | 360 |
| gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatcctc | 420 |
| accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag | 480 |
| gggtacctca tcaccccggc ctggctttgg gaaaagtacg gcctgaggcc cgaccagtgg | 540 |
| gccgactacc gggccctgac cggggacgag tccgacaacc ttcccgggggt caagggcatc | 600 |

aagatcgcca tgggtcaagct cttccccagg ctaaagcccc tgggggcccc cctcctcctc 2340
 caagtggcca acgagctgggt cctggagggtg cccgaggacc gggccgagga ggccaaggcc 2400
 ctggtcaagg aggtcatgga gaacgcctac cccctggacg tgccccctcga ggtggagggtg 2460
 ggcggtgggtc gggactgggt ggaggcgaag caggat 2496

<210> 446
 <211> 832
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 446

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Asn | Ser | Gly | Met | Leu | Pro | Leu | Phe | Glu | Pro | Lys | Gly | Arg | Val | Leu |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Leu | Val | Asp | Gly | His | His | Leu | Ala | Tyr | Arg | Thr | Phe | His | Ala | Leu | Lys |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Gly | Leu | Thr | Thr | Ser | Arg | Gly | Glu | Pro | Val | Gln | Ala | Val | Tyr | Gly | Phe |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Ala | Lys | Ser | Leu | Leu | Lys | Ala | Leu | Lys | Glu | Asp | Gly | Asp | Ala | Val | Ile |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| Val | Val | Phe | Asp | Ala | Lys | Ala | Pro | Ser | Phe | Arg | His | Glu | Ala | Tyr | Gly |
| 65 | | | | | 70 | | | | | 75 | | | | | 80 |
| Gly | Tyr | Lys | Ala | Gly | Arg | Ala | Pro | Thr | Pro | Glu | Asp | Phe | Pro | Arg | Gln |
| | | | | 85 | | | | | 90 | | | | | 95 | |
| Leu | Ala | Leu | Ile | Lys | Glu | Leu | Val | Asp | Leu | Leu | Gly | Phe | Thr | Arg | Leu |
| | | | 100 | | | | | 105 | | | | | 110 | | |
| Glu | Val | Pro | Gly | Tyr | Glu | Ala | Asp | Asp | Val | Leu | Ala | Ser | Leu | Ala | Lys |
| | | 115 | | | | | 120 | | | | | 125 | | | |
| Lys | Ala | Glu | Lys | Glu | Gly | Tyr | Glu | Val | Arg | Ile | Leu | Thr | Ala | Asp | Lys |
| | | 130 | | | | 135 | | | | | 140 | | | | |
| Asp | Leu | Tyr | Gln | Leu | Leu | Ser | Asp | Arg | Ile | His | Val | Leu | His | Pro | Glu |
| 145 | | | | | 150 | | | | | 155 | | | | | 160 |
| Gly | Tyr | Leu | Ile | Thr | Pro | Ala | Trp | Leu | Trp | Glu | Lys | Tyr | Gly | Leu | Arg |
| | | | | 165 | | | | | 170 | | | | | 175 | |
| Pro | Asp | Gln | Trp | Ala | Asp | Tyr | Arg | Ala | Leu | Thr | Gly | Asp | Glu | Ser | Asp |
| | | | 180 | | | | | 185 | | | | | 190 | | |
| Asn | Leu | Pro | Gly | Val | Lys | Gly | Ile | Gly | Glu | Lys | Thr | Ala | Arg | Lys | Leu |
| | | 195 | | | | | 200 | | | | | 205 | | | |

Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
 210 215 220
 Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
 225 230 235 240
 Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
 245 250 255
 Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
 260 265 270
 Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
 275 280 285
 Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu
 290 295 300
 Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala
 305 310 315 320
 Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala
 325 330 335
 Thr Ser Pro Val Glu Ala Leu Ala Asp Leu Lys Glu Ala Arg Gly Phe
 340 345 350
 Leu Ala Lys Asp Leu Ala Val Leu Ala Leu Arg Glu Gly Val Ala Leu
 355 360 365
 Asp Pro Thr Asp Asp Pro Leu Leu Val Ala Tyr Leu Leu Asp Pro Ala
 370 375 380
 Asn Thr His Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Phe Thr
 385 390 395 400
 Glu Asp Ala Ala Glu Arg Ala Leu Leu Ser Glu Arg Leu Phe Gln Asn
 405 410 415
 Leu Phe Lys Arg Leu Ser Glu Lys Leu Leu Trp Leu Tyr Gln Glu Val
 420 425 430
 Glu Arg Pro Leu Ser Arg Val Leu Ala His Met Glu Ala Arg Gly Val
 435 440 445
 Arg Leu Asp Val Pro Leu Leu Glu Ala Leu Ser Phe Glu Leu Glu Lys
 450 455 460
 Glu Met Glu Arg Leu Glu Gly Glu Val Phe Arg Leu Ala Gly His Pro
 465 470 475 480
 Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe Asp Glu
 485 490 495
 Leu Gly Leu Thr Pro Val Lys Arg Thr Lys Lys Thr Gly Lys Arg Ser
 500 505 510

| | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Thr | Ala | Gln | Gly | Ala | Leu | Glu | Ala | Leu | Arg | Gly | Ala | His | Pro | Ile | Val | 515 | 520 | 525 |
| Glu | Leu | Ile | Leu | Gln | Tyr | Arg | Glu | Leu | Ser | Lys | Leu | Lys | Ser | Thr | Tyr | 530 | 535 | 540 |
| Leu | Asp | Pro | Leu | Pro | Arg | Leu | Val | His | Pro | Arg | Thr | Gly | Arg | Leu | His | 545 | 550 | 555 |
| Thr | Arg | Phe | Asn | Gln | Thr | Ala | Thr | Ala | Thr | Gly | Arg | Leu | Ser | Ser | Ser | 565 | 570 | 575 |
| Asp | Pro | Asn | Leu | Gln | Asn | Ile | Pro | Val | Arg | Thr | Pro | Leu | Gly | Gln | Arg | 580 | 585 | 590 |
| Ile | Arg | Lys | Ala | Phe | Val | Ala | Glu | Glu | Gly | Trp | Leu | Leu | Leu | Ala | Ala | 595 | 600 | 605 |
| Asp | Tyr | Ser | Gln | Ile | Glu | Leu | Arg | Val | Leu | Ala | His | Leu | Ser | Gly | Asp | 610 | 615 | 620 |
| Glu | Asn | Leu | Lys | Arg | Val | Phe | Arg | Glu | Gly | Lys | Asp | Ile | His | Thr | Glu | 625 | 630 | 635 |
| Thr | Ala | Ala | Trp | Met | Phe | Gly | Leu | Asp | Pro | Ala | Leu | Val | Asp | Pro | Lys | 645 | 650 | 655 |
| Met | Arg | Arg | Ala | Ala | Lys | Thr | Val | Asn | Phe | Gly | Val | Leu | Tyr | Gly | Met | 660 | 665 | 670 |
| Ser | Ala | His | Arg | Leu | Ser | Gln | Glu | Leu | Gly | Ile | Asp | Tyr | Lys | Glu | Ala | 675 | 680 | 685 |
| Glu | Ala | Phe | Ile | Glu | Arg | Tyr | Phe | Gln | Ser | Phe | Pro | Lys | Val | Arg | Ala | 690 | 695 | 700 |
| Trp | Ile | Glu | Arg | Thr | Leu | Glu | Glu | Gly | Arg | Thr | Arg | Gly | Tyr | Val | Glu | 705 | 710 | 715 |
| Thr | Leu | Phe | Gly | Arg | Arg | Arg | Tyr | Val | Pro | Asp | Leu | Ala | Ser | Arg | Val | 725 | 730 | 735 |
| Arg | Ser | Val | Arg | Glu | Ala | Ala | Glu | Arg | Met | Ala | Phe | Asn | Met | Pro | Val | 740 | 745 | 750 |
| Gln | Gly | Thr | Ala | Ala | Asp | Leu | Met | Lys | Ile | Ala | Met | Val | Lys | Leu | Phe | 755 | 760 | 765 |
| Pro | Arg | Leu | Lys | Pro | Leu | Gly | Ala | His | Leu | Leu | Leu | Gln | Val | Ala | Asn | 770 | 775 | 780 |
| Glu | Leu | Val | Leu | Glu | Val | Pro | Glu | Asp | Arg | Ala | Glu | Glu | Ala | Lys | Ala | 785 | 790 | 795 |
| Leu | Val | Lys | Glu | Val | Met | Glu | Asn | Ala | Tyr | Pro | Leu | Asp | Val | Pro | Leu | 805 | 810 | 815 |

Glu Val Glu Val Gly Val Gly Arg Asp Trp Leu Glu Ala Lys Gln Asp
820 825 830

<210> 447
<211> 2508
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 447
atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
gaaccggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac 180
gggtacaagg ccgtcttcgt ggtctttgac gccaaaggccc cctccttccg ccacgaggcc 240
tacgaggcct acaaggcggg gagggccccg acccccaggg acttcccccg gcagctcgcc 300
ctcatcaagg agctggtgga cctcctgggg ttaccgccg tcgagggtccc cggtctacgag 360
gcggacgacg ttctcgccac cctggccaag aaggcgaaa aggaggggta cgaggtgcgc 420
atcctcaccg ccgaccgga cctctaccaa ctctctccg accgcgtcgc cgtcctccac 480
cccgagggcc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag 540
cagtgggtgg acttcgcgc cctcgtgggg gaccctccg acaacctccc cgggggtcaag 600
ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc 660
ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg 720
gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctccc cctggagggtg 780
gacctcgccc aggggcggga gcccgaccgg gaggggctta gggccttcct ggagaggctg 840
gagttcggca gcctcctcca cgagttcggc ctcttgagg ccccgcccc cctggaggag 900
gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctcccg ccccgagccc 960
atgtgggcgg agcttaaagc cctggccgcc tgcaggggcg gccgcgttta ccgggcggag 1020
gatcccttgg aggccttgcg ggggcttggg gaggtgaggg ggcttttggc caaggacctg 1080
gcgggtgctg ccctgaggga agggattgcc ctggcaccgg gcgacgacct catgctcctc 1140
gcctacctcc tggatccttc caacaccgcc cccgaagggg tagcccgggc ctacgggggg 1200
gagtggaccg aggaggcggg ggaaaggcg ctgctttccg aaaggcttta cgccgccctc 1260
ctgaagcggc ttaaggggga ggagaggctt ctttggcttt acgaggaggt ggaaaagccc 1320

0506436-05404

```

ctttcgcggg tcctggccca catggaggcc acgggggtac ggttgatgt ggcctactta 1380
aaggcccttt ccctggaggt ggaggcggag ataaggcgct tcgaggagga ggtccaccgc 1440
ctggccgggc atcctttcaa cctgaactcc cgggaccagc tggaaagggt catctttgac 1500
gagcttgggc ttcccgccat caagaagacg aggaagacgg gcaagcgctc caccagcgcc 1560
gccgttttgg aggccttgcg ggaggctcat cccatcgtag accgcatcct tcagtaccgg 1620
gagctttcca agctcaaggg aacctacatc gatcccttgc ctgccctggg ccaccccaag 1680
acgaaccgcc tccacaccgc tttcaaccag acggccaccg ccacggggag gcttagcagc 1740
tcggatccta atctgcaaaa tatccccgtg cgcaccctt tgggccagcg gatccgccgg 1800
gccttcgtgg ccgaggaggg gtggaggctg gtggttttgg actacagcca gattgagctc 1860
agggtcctgg cgcacctttc cggggacgag aacctaattc gggctttcca ggagggccag 1920
gacatccaca ccagacggc cagctggatg ttccggctgc cccagaggc cgtggattcc 1980
ctgatgcgcc gggcgcccaa gaccatcaac ttccggctcc tctacggcat gtccgcccac 2040
cggctttcgg gagagctggc catcccctac gaggaggcgg tggccttcat cgagcgggat 2100
ttccagagct accccaaggt gcgggcctgg attgagaaaa ccctggcgga aggacgggaa 2160
cggggctatg tggaaaccct ctttggccgc cggcgctacg tgcccgaact ggcttcccgg 2220
gtgaagagca tccgggaggg agcggagcgc atggccttca acatgccggg ccaggggacc 2280
gccgcggatt tgatgaaact ggccatggtg aagctctttc ccaggcttca ggagctgggg 2340
gccaggatgc ttttgcaggt gcacaacgaa ctggctctcg aggctcccaa ggagcaagcg 2400
gaggaagtgc cccaggaggc caagcggacc atggaggagg tgtggcccct gaaggtgccc 2460
ttggaggtgg aagtgggcat cggggaggac tggctttccg ccaaggcc 2508

```

<210> 448
 <211> 836
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 448

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Asn | Ser | Glu | Ala | Met | Leu | Pro | Leu | Phe | Glu | Pro | Lys | Gly | Arg | Val |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Leu | Leu | Val | Asp | Gly | His | His | Leu | Ala | Tyr | Arg | Thr | Phe | Phe | Ala | Leu |
| | | | 20 | | | | | 25 | | | | | | 30 | |

| | |
|---|------|
| ctcatcaagg agctggtgga cctcctgggg tttacccgcc tcgaggtccc cggctacgag | 360 |
| geggacgacg ttctcgccac cctggccaag aaggcggaaa aggaggggta cgaggtgcgc | 420 |
| atcctcaccg ccgaccgga cctctaccaa ctctgtctccg accgcgtcgc cgtcctccac | 480 |
| cccaggggcc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag | 540 |
| cagtgggtgg acttccgcgc cctcgtgggg gaccctccg acaacctccc cggggtcaag | 600 |
| ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc | 660 |
| ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg | 720 |
| gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctccc cctggaggtg | 780 |
| gacctcgcgc aggggcggga gcccgaccgg gaggggctta gggccttctt ggagaggctg | 840 |
| gagttcggca gcctcctcca cgagttcggc ctcttgagg ccccgcccc cctggaggag | 900 |
| gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctcccg ccccgagccc | 960 |
| atgtgggcgg agcttaaagc cctggccgcc tgcaggggcg gccgcgtcca ccgggcaaca | 1020 |
| agcccgttg aggccctggc cgacctcaag gagggccggg ggttctctggc caaggacctg | 1080 |
| gccgttttgg ccctgcggga gggggtggcc ctggaccca ccgacgacct cctcctggtg | 1140 |
| gcctacctcc tggaccggc caacaccac cccgaggggg tggcccggcg ctacgggggc | 1200 |
| gagttcacgg aggacgcagc ggagagggcc ctctctccg agaggctctt ccagaacctc | 1260 |
| tttaaacggc tttccgagaa gtcctctgg ctctaccagg aggtggagcg gcccctctcc | 1320 |
| cgggtcttgg ccacatgga ggcccgggg gtgaggctgg acgtccccct tctggaggcc | 1380 |
| ctctcctttg agctggagaa ggagatggag cgcctggagg gggaggtctt ccgtttggcc | 1440 |
| ggccacctct tcaacctcaa ctcccgcgac cagctggaaa gggctctctt tgacgagctg | 1500 |
| ggcctcacc cgggtgaagc gacgaagaag acgggcaagc gctccaccgc ccagggggcc | 1560 |
| ctggaggccc tccggggggc ccacccatc gtggagctca tcctccagta ccgggagctt | 1620 |
| tccaagctca aaagcaccta cctggacccc ctgccccggc tcgtccacct gcggacgggc | 1680 |
| cggctccaca cccgcttcaa ccagacggcc acggccacgg gaaggcttct cagctccgac | 1740 |
| cccaacctgc agaacatccc cgtgcgaccc cccttggggc agcgcatccg caaggccttc | 1800 |
| gtggccgagg aggggtggct ctttttggcg gcggactact ccagattga gctccgggtc | 1860 |
| ctggcccacc tctcggggga cgagaacctg aagcgggtct tccgggaggg gaaggacatc | 1920 |
| cataccgaga ccgcgcctg gatgttcggc ttagacccc ctctggtgga tccaaagatg | 1980 |

cgccgggcg ccaagacggt caacttcggc gtccctctacg ggatgtccgc ccacaggctc 2040
 tcccaggagc tcggcataga ctacaaggag gcgagggcct ttattgagcg ctacttccag 2100
 agcttcccca aggtgcgggc ctggatagaa aggaccctgg aggagggccg gacgcggggc 2160
 tacgtggaga ccctgttcgg caggaggcgc tatgtgcccg acctggcctc ccgggtccgc 2220
 tcggtgcggg aggcggcgga gcggatggcc ttcaacatgc ccgtgcaggg caccgccgcc 2280
 gacctgatga agatcgccat ggtcaagctc ttccccaggc taaagcccct gggggccac 2340
 ctctctctcc aagtggccaa cgagctggc ctggaggtgc ccgaggaccg ggccgaggag 2400
 gccaaagccc tggtaagga ggtcatggag aacgcctacc ccctggacgt gcccctcgag 2460
 gtggaggtgg gcgtgggtcg ggactggctg gaggcgaagc aggat 2505

<210> 450
 <211> 835
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 450

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Asn | Ser | Glu | Ala | Met | Leu | Pro | Leu | Phe | Glu | Pro | Lys | Gly | Arg | Val |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Leu | Leu | Val | Asp | Gly | His | His | Leu | Ala | Tyr | Arg | Thr | Phe | Phe | Ala | Leu |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Lys | Gly | Leu | Thr | Thr | Ser | Arg | Gly | Glu | Pro | Val | Gln | Ala | Val | Tyr | Gly |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Phe | Ala | Lys | Ser | Leu | Leu | Lys | Ala | Leu | Lys | Glu | Asp | Gly | Tyr | Lys | Ala |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| Val | Phe | Val | Val | Phe | Asp | Ala | Lys | Ala | Pro | Ser | Phe | Arg | His | Glu | Ala |
| 65 | | | | | 70 | | | | | 75 | | | | 80 | |
| Tyr | Glu | Ala | Tyr | Lys | Ala | Gly | Arg | Ala | Pro | Thr | Pro | Glu | Asp | Phe | Pro |
| | | | | 85 | | | | | 90 | | | | | 95 | |
| Arg | Gln | Leu | Ala | Leu | Ile | Lys | Glu | Leu | Val | Asp | Leu | Leu | Gly | Phe | Thr |
| | | 100 | | | | | 105 | | | | | | 110 | | |
| Arg | Leu | Glu | Val | Pro | Gly | Tyr | Glu | Ala | Asp | Asp | Val | Leu | Ala | Thr | Leu |
| | | 115 | | | | | 120 | | | | | 125 | | | |
| Ala | Lys | Lys | Ala | Glu | Lys | Glu | Gly | Tyr | Glu | Val | Arg | Ile | Leu | Thr | Ala |
| | 130 | | | | | 135 | | | | | 140 | | | | |
| Asp | Arg | Asp | Leu | Tyr | Gln | Leu | Val | Ser | Asp | Arg | Val | Ala | Val | Leu | His |

| | | | |
|--|-----|-----|-----|
| 755 | 760 | 765 | |
| Lys Leu Phe Pro Arg Leu Lys Pro Leu Gly Ala His Leu Leu Leu Gln | | | |
| 770 | 775 | 780 | |
| Val Ala Asn Glu Leu Val Leu Glu Val Pro Glu Asp Arg Ala Glu Glu | | | |
| 785 | 790 | 795 | 800 |
| Ala Lys Ala Leu Val Lys Glu Val Met Glu Asn Ala Tyr Pro Leu Asp | | | |
| | 805 | 810 | 815 |
| Val Pro Leu Glu Val Glu Val Gly Val Gly Arg Asp Trp Leu Glu Ala | | | |
| | 820 | 825 | 830 |
| Lys Gln Asp | | | |
| 835 | | | |
| <210> 451 | | | |
| <211> 24 | | | |
| <212> DNA | | | |
| <213> Artificial | | | |
| <220> | | | |
| <223> Synthetic | | | |
| <400> 451 | | | |
| cttctctcat ccgccaaaac agcc | | | 24 |
| <210> 452 | | | |
| <211> 2526 | | | |
| <212> DNA | | | |
| <213> Artificial | | | |
| <220> | | | |
| <223> Synthetic | | | |
| <400> 452 | | | |
| atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctgggtggac | | | 60 |
| ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc | | | 120 |
| gaaccggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggcctt gaaggaggac | | | 180 |
| gggtacaagg ccgtcatcgt ggtctttgac gccgaggccc cctccttcctg ccacgaggcc | | | 240 |
| tacgaggcct acaaggcggg gagggccccg acccccagagg acttcccccg gcagctcgcc | | | 300 |
| ctcatcaagg agctggtgga cctcctgggg tttaccgcgc tcgaggtccc cggctacgag | | | 360 |
| gcggacgacg ttctcgccac cctggccaag aaggcggaaa aggaggggta cgaggtgcgc | | | 420 |
| atcctcaccg ccgaccgca cctctaccaa ctcgctctccg accgcgtcgc cgtcctccac | | | 480 |
| cccaggggcc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag | | | 540 |
| cagtgggtgg acttccgcgc cctcgtgggg gaccctccg acaacctccc cggggtcaag | | | 600 |

| | | | | | | |
|------------|-------------|-------------|-------------|------------|-------------|------|
| cgccgggagc | cggaccggga | ggggcttaag | gccttttttg | agaggctgga | gttcggaagc | 840 |
| ctcctccacg | agttcggcct | gttggaagc | ccggtggcgg | cggaggaagc | tccctggccg | 900 |
| cccccgagg | gagccttcgt | ggggtacgtt | ctttcccgcc | ccgagcccat | gtgggcggag | 960 |
| cttaacgcct | tggccgcgcg | ctggggcggc | cgcgtagcacc | gggcagcaga | ccccttggcg | 1020 |
| gggctaaagg | acctcaagga | ggtccggggc | ctcctcgcca | aggacctcgc | cgtcttggcc | 1080 |
| tcgagggagg | ggctagacct | cgtgcccggg | gacgacccca | tgctcctcgc | ctacctcctg | 1140 |
| gaccttctga | acaccacccc | cgaggggggtg | gcgcggcgct | acggggggga | gtggacggag | 1200 |
| gacgcgcgcc | accgggcctt | cctctcggag | aggctccatc | ggaacctcct | taagcgcctc | 1260 |
| gagggggagg | agaagctcct | ttggctctac | cacgaggtgg | aaaagcccct | ctcccgggtc | 1320 |
| ctggcccata | tggaggccac | cggggtacgg | cgggacgtgg | cctaccttca | ggccctttcc | 1380 |
| ctggagcttg | cggaggagat | ccgccgcctc | gaggaggagg | tcttcgcgtt | ggcggggccac | 1440 |
| cccttcaacc | tcaactcccg | ggaccagctg | gaaaggggtgc | tctttgacga | gcttaggctt | 1500 |
| cccgccttga | agaagacgaa | gaagacaggc | aagcgctcca | ccagcgccgc | ggtgctggag | 1560 |
| gccctacggg | aggcccaccc | catcgtggag | aagatcctcc | agcaccggga | gctcaccaag | 1620 |
| ctcaagaaca | cctacgtgga | ccccctccca | agcctcgtcc | acccgaggac | gggcgcgcctc | 1680 |
| cacacccgct | tcaaccagac | ggccacggcc | acggggaggc | ttagtagctc | cgaccccaac | 1740 |
| ctgcagaaca | tccccgtccg | cacccccttg | ggccagagga | tccgcccggc | cttcgtggcc | 1800 |
| gaggcggggt | gggcgttggt | ggccctggac | tatagccaga | tagagctccg | cgtcctcgcc | 1860 |
| cacctctccg | gggacgaaaa | cctgatcagg | gtcttccagg | aggggaagga | catccacacc | 1920 |
| cagaccgcaa | gctggatggt | cggcgtcccc | ccggaggccg | tggaccccct | gatgcgcggg | 1980 |
| gcggccaaga | cggatgaactt | cggcgtcctc | tacggcatgt | ccgcccatag | gctctcccag | 2040 |
| gagcttgcca | tcccctacga | ggaggcggtg | gcctttatag | agcgtactt | ccaaagcttc | 2100 |
| cccaagggtg | gggcctggat | agaaaagacc | ctggaggagg | ggaggaagcg | gggctacgtg | 2160 |
| gaaaccctct | tcggaagaag | gcgctacgtg | cccgaacctca | acgcccggtg | gaagagcgtc | 2220 |
| agggaggccg | cggagcgcat | ggccttcaac | atgcccgctc | agggcaccgc | cgccgacctc | 2280 |
| atgaagctcg | ccatggtgaa | gctcttcccc | cgcctccggg | agatgggggc | ccgcatgctc | 2340 |
| ctccaggtcg | ccaacgagct | cctcctggag | gcccccaag | cgcgggccga | ggaggtggcg | 2400 |
| gctttggcca | aggaggccat | ggagaaggcc | tatccccctg | ccgtgccctt | ggaggtggag | 2460 |

gtggggatgg gggaggactg gcttttcgcc aagggtcacc accaccacca ccac 2514

<210> 455
<211> 2526
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 455
atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
gaaccggtgc aggcggtcta cggtctcgcc aagagcctcc tcaaggccct gaaggaggac 180
gggtacaagg ccgtcatcgt ggtctttgac gccgaggccc cctccttcg ccacgaggcc 240
tacgaggcct acaaggcggg gagggccccg acccccgagg acttccccg gcagctcgcc 300
ctcatcaagg agctggtgga cctcctgggg tttaccgcc tcgagggtccc cggctacgag 360
gcgacgacg ttctcgccac cctggccaag aaggcggaaa aggaggggta cgaggtgcgc 420
atcctcaccg ccgaccgga cctctaccaa ctgctctccg accgctcgc cgtcctccac 480
cccgagggcc acctcatcac cccggagtgg ctttgggaga agtatgggct taagccttcc 540
cagtgggtgg actaccgggc cttggccggg gacccttcg acaacatccc cggcgtgaag 600
ggcatcgggg agaagacggc ggccaagctg atccgggagt ggggaagcct ggaaaacctt 660
cttaagcacc tggaacaggt gaaacctgcc tccgtgcggg agaagatcct tagccacatg 720
gaggacctca agctatccct ggagctatcc cgggtgcaca cggacttgct ccttcagggtg 780
gacttcgccc ggccgggga gccggaccgg gaggggctta aggccttttt ggagaggetg 840
gagttcgaa gcctcctcca cgagttcggc ctgttggaag gcccggtggc ggccggaggaa 900
gtccctggc cggccccga gggagccttc gtggggtacg ttctttcccg ccccgagccc 960
atgtgggagg agcttaacgc cttggccgcc gcctggggcg gccgcgtgca ccgggcagca 1020
gacctcttgg cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc 1080
gccgtcttgg cctcgaggga ggggctagac ctgctgcccg gggacgacct catgctcctc 1140
gcctacctcc tggaccttc gaacaccacc cccgaggggg tggcgcggcg ctacgggggg 1200
gagtggacgg aggacgccgc ccaccgggcc ctctctcgg agaggctcca tcggaacctc 1260
cttaagcgcc tcgaggggga ggagaagctc ctttggctct accacgaggt ggaaaagccc 1320
ctctccggg tcctggccca tatggaggcc accggggtac ggcgggacgt ggccctacctt 1380

Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
 35 40 45
 Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala
 50 55 60
 Val Ile Val Val Phe Asp Ala Glu Ala Pro Ser Phe Arg His Glu Ala
 65 70 75 80
 Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro
 85 90 95
 Arg Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr
 100 105 110
 Arg Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu
 115 120 125
 Ala Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala
 130 135 140
 Asp Arg Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His
 145 150 155 160
 Pro Glu Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly
 165 170 175
 Leu Lys Pro Ser Gln Trp Val Asp Tyr Arg Ala Leu Ala Gly Asp Pro
 180 185 190
 Ser Asp Asn Ile Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Ala
 195 200 205
 Lys Leu Ile Arg Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys His Leu
 210 215 220
 Glu Gln Val Lys Pro Ala Ser Val Arg Glu Lys Ile Leu Ser His Met
 225 230 235 240
 Glu Asp Leu Lys Leu Ser Leu Glu Leu Ser Arg Val His Thr Asp Leu
 245 250 255
 Leu Leu Gln Val Asp Phe Ala Arg Arg Arg Glu Pro Asp Arg Glu Gly
 260 265 270
 Leu Lys Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu
 275 280 285
 Phe Gly Leu Leu Glu Ser Pro Val Ala Ala Glu Glu Ala Pro Trp Pro
 290 295 300
 Pro Pro Glu Gly Ala Phe Val Gly Tyr Val Leu Ser Arg Pro Glu Pro
 305 310 315 320
 Met Trp Ala Glu Leu Asn Ala Leu Ala Ala Ala Trp Gly Gly Arg Val
 325 330 335

Asp Ile His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu
645 650 655

Ala Val Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly
660 665 670

Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile
675 680 685

Pro Tyr Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe
690 695 700

Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys
705 710 715 720

Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp
725 730 735

Leu Asn Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala
740 745 750

Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala
755 760 765

Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu
770 775 780

Leu Gln Val Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala
785 790 795 800

Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro
805 810 815

Leu Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu
820 825 830

Ser Ala Lys Gly His His His His His His
835 840

<210> 457
<211> 27
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 457
ttccaggtgc ttgaggaggt tttccag

27

<210> 458
<211> 27
<212> DNA
<213> Artificial

<220>

<223> Synthetic

<400> 458

ctcctcaagc acctggaaca ggtgaaa

27

<210> 459

<211> 2526

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 459

atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
gaaccgggtgc aggcgggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac 180
gggtacaagg ccgtcatcgt ggtctttgac gccgaggccc cctccttcg ccacgaggcc 240
tacgaggcct acaaggcggg gagggccccg acccccgagg acttcccccg gcagctcgcc 300
ctcatcaagg agctggtgga cctcctgggg tttaccgcc tcgaggtccc cggctacgag 360
gcggacgacg ttctcgccac cctggccaag aaggcggaag aggaggggta cgaggtgcgc 420
atcctcaccg ccgaccgca cctctaccaa ctcgctctcg accgcgtcgc cgtcctccac 480
cccgagggcc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag 540
cagtgggtgg acttccgcgc cctcgtgggg gaccctccg acaacctccc cgggggtcaag 600
ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc 660
ctcaagcacc tggaacaggt gaaacctgcc tccgtgcggg agaagatcct tagccacatg 720
gaggacctca agctatccct ggagctatcc cgggtgcaca cggacttgct ccttcagggtg 780
gacttcgccc ggccggggga gccggaccgg gaggggctta aggccttttt ggagaggctg 840
gagttcggaa gcctcctcca cgagttcggc ctggttgaaa gcccggtggc ggccggaggaa 900
gtccctggc cgcgggggga gggagccttc gtggggtacg ttctttcccg ccccgagccc 960
atgtgggcgg agcttaacgc cttggccgcc gcctggggcg gcccggtgca ccgggcagca 1020
gacccttggt cgggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc 1080
gccgtcttgg cctcgaggga ggggctagac ctcggtcccc gggacgaccc catgctcctc 1140
gcctacctcc tggacccttc gaacaccacc cccgaggggg tggcgcggcg ctacgggggg 1200
gagtggacgg aggaagccgc ccaccgggcc ctcctctcgg agaggctcca tcggaacctc 1260

| | | | |
|---|-----|-----|-----|
| 1 | 5 | 10 | 15 |
| Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu | 20 | 25 | 30 |
| Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly | 35 | 40 | 45 |
| Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala | 50 | 55 | 60 |
| Val Ile Val Val Phe Asp Ala Glu Ala Pro Ser Phe Arg His Glu Ala | 65 | 70 | 75 |
| Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro | 85 | 90 | 95 |
| Arg Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr | 100 | 105 | 110 |
| Arg Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu | 115 | 120 | 125 |
| Ala Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala | 130 | 135 | 140 |
| Asp Arg Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His | 145 | 150 | 155 |
| Pro Glu Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly | 165 | 170 | 175 |
| Leu Arg Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro | 180 | 185 | 190 |
| Ser Asp Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu | 195 | 200 | 205 |
| Lys Leu Leu Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys His Leu | 210 | 215 | 220 |
| Glu Gln Val Lys Pro Ala Ser Val Arg Glu Lys Ile Leu Ser His Met | 225 | 230 | 235 |
| Glu Asp Leu Lys Leu Ser Leu Glu Leu Ser Arg Val His Thr Asp Leu | 245 | 250 | 255 |
| Leu Leu Gln Val Asp Phe Ala Arg Arg Arg Glu Pro Asp Arg Glu Gly | 260 | 265 | 270 |
| Leu Lys Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu | 275 | 280 | 285 |
| Phe Gly Leu Leu Glu Ser Pro Val Ala Ala Glu Glu Ala Pro Trp Pro | 290 | 295 | 300 |
| Pro Pro Glu Gly Ala Phe Val Gly Tyr Val Leu Ser Arg Pro Glu Pro | | | |

| | | |
|---|-----|-------------|
| 610 | 615 | 620 |
| His Leu Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys | | |
| 625 | 630 | 635 640 |
| Asp Ile His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu | | |
| | 645 | 650 655 |
| Ala Val Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly | | |
| | 660 | 665 670 |
| Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile | | |
| | 675 | 680 685 |
| Pro Tyr Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe | | |
| | 690 | 695 700 |
| Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys | | |
| | 705 | 710 715 720 |
| Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp | | |
| | 725 | 730 735 |
| Leu Asn Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala | | |
| | 740 | 745 750 |
| Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala | | |
| | 755 | 760 765 |
| Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu | | |
| | 770 | 775 780 |
| Leu Gln Val Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala | | |
| | 785 | 790 795 800 |
| Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro | | |
| | 805 | 810 815 |
| Leu Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu | | |
| | 820 | 825 830 |
| Ser Ala Lys Gly His His His His His His | | |
| | 835 | 840 |
| <210> 461 | | |
| <211> 27 | | |
| <212> DNA | | |
| <213> Artificial | | |
| <220> | | |
| <223> Synthetic | | |
| <400> 461 | | |
| gtccaggttc ttgaggaggt tttccag | | |

<211> 27
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 462
 ctctcaaga acctggaccg ggtaaag

27

<210> 463
 <211> 2532
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 463
 atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
 ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
 gaaccggtgc aggcggtcta cggtctcgcc aagagcctcc tcaaggccct gaaggaggac 180
 ggggtacaagg ccgtcatcgt ggtctttgac gccgaggccc cctccttccg ccacgaggcc 240
 tacgaggcct acaaggcggg gagggccccg acccccagagg acttcccccg gcagctcgcc 300
 ctcatcaagg agctggtgga cctcctgggg tttacccgcc tcgagggtccc cggctacgag 360
 gcgagcgacg ttctcgccac cctggccaag aaggcggaag aggaggggta cgagggtgcgc 420
 atcctcaccg ccgaccgga cctctaccaa ctgctctccg accgcgtcgc cgtcctccac 480
 cccgagggcc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag 540
 cagtgggtgg acttccgcgc cctcgtgggg gacctctccg acaacctccc cggggtcaag 600
 ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc 660
 ctcaagaacc tggaccgggt aaagccggac tcgctccggc gcaagataga ggcgcacctc 720
 gaggacctcc acctctcctt agacctggcc cgcacccgca ccgacctccc cctggagggtg 780
 gactttaagg cctgcgcgc caggaccccc gacctggagg gcctgagggc ctttttggag 840
 gagctggagt tcggaagcct cctccacgag ttcggcctcc tgggagggga gaagccccgg 900
 gaggaggccc cctggccccc gcccggaagg gccttcgtgg gcttctcctt tccccgcaag 960
 gagcccatgt gggcgagct tctggccctg gcggcggcct cgggcggccg cgtgcaccgg 1020
 gcagcagacc ccttggcggg gctaaaggac ctcaaggagg tccggggcct cctcgccaag 1080
 gacctcgccg tcttggcctc gagggagggg ctagacctcg tgcccgggga cgaccccatg 1140

ctcctcgctt acctcctgga cccttcgaac accacccccg aggggggtggc gcggcgctac 1200
 ggggggggagt ggacggagga cgccgcccac cgggcccctcc tctcggagag gctccatcgg 1260
 aacctcctta agcgccctga gggggaggag aagctccttt ggctctacca cgaggtggaa 1320
 aagccccctct cccgggtcct ggcccatatg gagggcaccg gggtagggcg ggacgtggcc 1380
 taccttcagg ccctttccct ggagcttgcg gaggagatcc gccgcctcga ggaggaggtc 1440
 ttccgcttgg cgggccaccc cttcaacctc aactcccggg accagctgga aagggtgctc 1500
 tttgacgagc ttaggcttcc cgccttgaag aagacgaaga agacaggcaa gcgctccacc 1560
 agcgccgcgg tgctggaggc cctacgggag gccaccccca tcgtggagaa gatcctccag 1620
 caccgggagc tcaccaagct caagaacacc tacgtggacc cctcccaag cctcgtccac 1680
 ccgaggacgg gccgcctcca caccgcttc aaccagacgg ccacggccac ggggaggctt 1740
 agtagctccg accccaacct gcagaacatc ccgctccgca ccccttggg ccagaggatc 1800
 cgccgggcct tcgtggccga ggcgggttgg gcgttggtgg ccctggacta tagccagata 1860
 gagctccgcg tcctcgccca cctctccggg gacgaaaacc tgatcagggt cttccaggag 1920
 gggaaggaca tccacaccca gaccgcaagc tggatgttcg gcgtccccc ggaggccgtg 1980
 gacccccctga tgcgccgggc ggccaagacg gtgaacttcg gcgtcctcta cggcatgtcc 2040
 gcccataggc tctcccagga gcttgccatc ccctacgagg aggcgggtggc ctttatagag 2100
 cgctacttcc aaagcttccc caagggtcgg gcttgatag aaaagaccct ggaggagggg 2160
 aggaagcggg gctacgtgga aacctcttc ggaagaaggc gctacgtgcc cgacctcaac 2220
 gcccggtgta agagcgtcag ggaggccgcg gagcgcatgg cttcaacat gccgtccag 2280
 ggcaccgccg ccgacctcat gaagctcgcc atggtgaagc tcttcccccg cctccgggag 2340
 atggggggccc gcatgctcct ccaggctgcc aacgagctcc tcctggaggc ccccaagcg 2400
 cgggccgagg aggtggcggc tttggccaag gaggccatgg agaaggccta tcccccgcc 2460
 gtgcccctgg aggtggaggt ggggatgggg gaggactggc tttccgcaa gggtcaccac 2520
 caccaccacc ac 2532

<210> 464
 <211> 844
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 464

Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
1 5 10 15
Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu
20 25 30
Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
35 40 45
Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala
50 55 60
Val Ile Val Val Phe Asp Ala Glu Ala Pro Ser Phe Arg His Glu Ala
65 70 75 80
Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro
85 90 95
Arg Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr
100 105 110
Arg Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu
115 120 125
Ala Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala
130 135 140
Asp Arg Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His
145 150 155 160
Pro Glu Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly
165 170 175
Leu Arg Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro
180 185 190
Ser Asp Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu
195 200 205
Lys Leu Leu Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu
210 215 220
Asp Arg Val Lys Pro Asp Ser Leu Arg Arg Lys Ile Glu Ala His Leu
225 230 235 240
Glu Asp Leu His Leu Ser Leu Asp Leu Ala Arg Ile Arg Thr Asp Leu
245 250 255
Pro Leu Glu Val Asp Phe Lys Ala Leu Arg Arg Arg Thr Pro Asp Leu
260 265 270
Glu Gly Leu Arg Ala Phe Leu Glu Glu Leu Glu Phe Gly Ser Leu Leu
275 280 285

gacgtccttc ggggtgatga ggtggcc 27

<210> 466
<211> 27
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 466
atcaccccgaggacgtcca ggagaag 27

<210> 467
<211> 2532
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 467
atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
gaaccggtgc aggcggtcta cggtctcgcc aagagcctcc tcaaggccct gaaggaggac 180
gggtacaagg ccgtcatcgt ggtctttgac gccgaggccc cctccttcg ccacgaggcc 240
tacgaggcct acaaggcggg gagggccccg acccccgagg acttccccg gcagctcgcc 300
ctcatcaagg agctggtgga cctcctgggg tttaccgcc tcgagggtccc cggctacgag 360
gcggacgacg ttctcgccac cctggccaag aaggcggaaa aggaggggta cgagggtgcg 420
atcctcaccg ccgaccgca cctctaccaa ctgctctccg accgctcgc cgtcctccac 480
cccgagggcc acctcatcac cccgaaggac gtccaggaga agtacggggt gcccccgag 540
cgctgggtgg acttccgcg cctcacgggg gaccgctcgg acaacatccc cgggggtggcg 600
gggatagggg agaagaccgc cttcgaactc ctgcagagt gggggagcgt ggaaaacctc 660
ctgaagaacc tggaccgggt aaagccggac tcgctccggc gcaagataga ggcgcacctc 720
gaggacctcc acctctcctt agacctggcc cgcctccgca ccgacctccc cctggagggtg 780
gactttaagg ccctgcgccg caggaccccc gacctggagg gcctgagggc ctttttgagg 840
gagctggagt tcggaagcct cctccacgag ttcggcctcc tgggagggga gaagccccgg 900
gaggaggccc cctggcccc ccccgaagg gccttcgtgg gcttctcct tccccgcaag 960
gagcccatgt gggcggagct tctggccctg gcggcgccct cgggcggccg cgtgcaccgg 1020

```

gcagcagacc ccttggcggg gctaaaggac ctcaaggagg tccggggcct cctcgccaag 1080
gacctcgccg tcttggcctc gagggagggg ctagacctcg tgcccgggga cgaccccatg 1140
ctcctcgccct acctcctgga cccttcgaac accacccccg aggggggtggc gcggcgctac 1200
gggggggaggt ggacggagga cgccgccac cgggccctcc tctcgagag gctccatcgg 1260
aacctcctta agcgctcga gggggaggag aagctccttt ggctctacca cgaggtggaa 1320
aagccccctct cccgggtcct ggcccatatg gagggcaccg gggtaggcg ggacgtggcc 1380
taccttcagg ccctttccct ggagcttgcg gaggagatcc gccgcctcga ggaggaggtc 1440
ttccgcttgg cggggcacc cttcaacctc aactcccggg accagctgga aagggtgctc 1500
tttgacgagc ttaggcttcc cgccttgaag aagacgaaga agacaggcaa gcgctccacc 1560
agcgccgcgg tgctggaggc cctacgggag gccacccca tcgtggagaa gatcctccag 1620
caccgggagc tcaccaagct caagaacacc tacgtggacc cctcccaag cctcgctccac 1680
ccgaggacgg gccgcctcca caccgcttc aaccagacgg ccacggccac ggggaggctt 1740
agtagctccg accccaacct gcagaacatc cccgtccgca ccccttggg ccagaggatc 1800
cgccgggcct tcgtggcga ggcggttgg gcgttgggg ccttgacta tagccagata 1860
gagctccgcg tctcgccta cctctccggg gacgaaaacc tgatcagggt cttccaggag 1920
gggaaggaca tccacacca gaccgaagc tggatgttcg gcgtccccc ggaggccgtg 1980
gacccccctga tgcgcggggc ggccaagacg gtgaacttcg gcgtcctcta cgcatgtcc 2040
gcccataggc tctcccagga gcttgccatc cctacgagg aggcggtggc ctttatagag 2100
cgctacttcc aaagcttccc caagggtcgg gcctggatag aaaagaccct ggaggagggg 2160
aggaagcggg gctacgtgga aaccctcttc ggaagaaggc gctacgtgcc cgacctaac 2220
ggccgggtga agagcgtcag ggaggccgag gagcgcatgg cttcaacat gcccgctcag 2280
ggcaccgcgg ccgacctcat gaagctcgcc atggtgaagc tcttccccg cctccgggag 2340
atggggggccc gcatgctcct ccaggtcgcc aacgagctcc tcttgaggc cccccaagcg 2400
cgggccgagg aggtggcggc ttggccaag gaggccatgg agaaggccta tcccctcgcc 2460
gtgcccctgg aggtggaggt ggggatgggg gaggactggc tttccgcaa gggtcaccac 2520
caccaccacc ac 2532

```

```

<210> 468
<211> 844
<212> PRT

```


Thr Gly Arg Leu Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val
580 585 590

Arg Thr Pro Leu Gly Gln Arg Ile Arg Arg Ala Phe Val Ala Glu Ala
595 600 605

Gly Trp Ala Leu Val Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val
610 615 620

Leu Ala His Leu Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu
625 630 635 640

Gly Lys Asp Ile His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro
645 650 655

Pro Glu Ala Val Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn
660 665 670

Phe Gly Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu
675 680 685

Ala Ile Pro Tyr Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln
690 695 700

Ser Phe Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly
705 710 715 720

Arg Lys Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val
725 730 735

Pro Asp Leu Asn Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg
740 745 750

Met Ala Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys
755 760 765

Leu Ala Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg
770 775 780

Met Leu Leu Gln Val Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala
785 790 795 800

Arg Ala Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala
805 810 815

Tyr Pro Leu Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp
820 825 830

Trp Leu Ser Ala Lys Gly His His His His His His
835 840

<210> 469
<211> 27
<212> DNA
<213> Artificial

<220>
 <223> Synthetic

<400> 469
 ctcgaggcgg gtaaacccca ggaggtc 27

<210> 470
 <211> 27
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 470
 gggtttacc gcctcgagg gcccggc 27

<210> 471
 <211> 2526
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 471
 atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
 ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
 gaaccggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac 180
 ggggtacaagg ccgtcatcgt ggtctttgac gccgaggccc cctccttccg ccacgaggcc 240
 tacgaggcct acaaggcggg gagggccccg acccccgagg acttcccccg gcagctcgcc 300
 ctcatcaagg agctggtgga cctcctgggg tttaccgcc tgcagggtgcc gggctttgaa 360
 gcggatgacg tcctggctac cctggccaag aaggcggaaa aggaaggcta cgaagtgcgc 420
 atcctcaccg cggaccggga cctttaccag cttctttcgg agcgaatctc catccttcac 480
 ccggagggtt acctgatcac cccggagtgg ctttgggaga agtatgggct taagccttcc 540
 cagtgggtgg actaccgggc cttggccggg gacccttccg acaacatccc cggcgtgaag 600
 ggcacgggg agaagacggc ggccaagctg atccgggagt ggggaagcct ggaaaacctt 660
 cttaagcacc tggaacaggt gaaacctgcc tccgtgcggg agaagatcct tagccacatg 720
 gaggacctca agctatccct ggagctatcc cgggtgcaca cggacttgct ccttcagggtg 780
 gacttcgccc ggcgccggga gccggaccgg gaggggctta aggccttttt ggagaggctg 840
 gagttcggaa gcctcctcca cgagttcggc ctgttggaag gcccgggtggc ggcggaggaa 900

| | |
|--|------|
| gtcccttggc cgccccccga gggagccttc gtggggtacg ttctttcccg ccccgagccc | 960 |
| atgtggggcg agcttaacgc cttggccgcc gcctggggcg gccgcgtgca ccgggcagca | 1020 |
| gaccccttgg cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc | 1080 |
| gccgtcttgg cctcgaggga ggggctagac ctcggtgccg gggacgaccc catgctcctc | 1140 |
| gcctacctcc tggaccttc gaacaccacc cccgaggggg tggcgcggcg ctacgggggg | 1200 |
| gagtggacgg aggacgccgc ccaccgggcc ctctctcgg agaggctcca tcggaacctc | 1260 |
| cttaagcgcc tcgaggggga ggagaagctc ctttggtctt accacgaggt ggaaaagccc | 1320 |
| ctctcccggg tcctggccca tatggaggcc accggggtac ggcgggacgt ggctacctt | 1380 |
| caggcccttt ccctggagct tgcggaggag atccgccgcc tcgaggagga ggtcttcgc | 1440 |
| ttggcggggc accccttcaa cctcaactcc cgggaccagc tggaaagggg gctctttgac | 1500 |
| gagcttaggc ttccgcctt gaagaagacg aagaagacag gcaagcgctc caccagcgcc | 1560 |
| gcggtgctgg agggcctacg ggaggccac cccatcgtgg agaagatcct ccagcaccgg | 1620 |
| gagctcacca agctcaagaa cacctacgtg gacccctcc caagcctcgt ccacccgagg | 1680 |
| acgggccgcc tccacaccg cttcaaccag acggccacgg ccacggggag gcttagtagc | 1740 |
| tccgaccca acctgcagaa catccccgtc cgcacccct tgggccagag gatccgccgg | 1800 |
| gccttcgtgg ccgaggcggg ttgggcgttg gtggccctgg actatagcca gatagagctc | 1860 |
| cgcgtcctcg cccacctctc cggggacgaa aacctgatca gggctctcca ggaggggaag | 1920 |
| gacatccaca cccagaccgc aagctggatg ttcggcgtcc ccccgaggc cgtggacccc | 1980 |
| ctgatgcgcc gggcggccaa gacggtgaac ttcggcgtcc tctacggcat gtccgcccac | 2040 |
| aggctctccc aggagcttgc catcccctac gaggaggcgg tggcctttat agagcgctac | 2100 |
| ttccaaagct tccccagggt gcgggccttg atagaaaaga ccctggagga ggggaggaag | 2160 |
| cggggctacg tggaaaccct cttcggaaga aggcgctacg tgcccgacct caacgcccgg | 2220 |
| gtgaagagcg tcagggaggc cgcgagcgc atggccttca acatgcccgt ccagggcacc | 2280 |
| gccgccgacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg | 2340 |
| gcccgcacgc tcctccaggt cgccaacgag ctctcctcgg agggccccca agcgcgggcc | 2400 |
| gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatcccct cgccgtgccc | 2460 |
| ctggaggtgg aggtggggat gggggaggac tggctttccg ccaaggggtca ccaccaccac | 2520 |
| caccac | 2526 |

| | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|-----|
| | | | | 245 | | | | | 250 | | | | | 255 | | | |
| Leu | Leu | Gln | Val | Asp | Phe | Ala | Arg | Arg | Arg | Glu | Pro | Asp | Arg | Glu | Gly | | |
| | | | 260 | | | | | 265 | | | | | 270 | | | | |
| Leu | Lys | Ala | Phe | Leu | Glu | Arg | Leu | Glu | Phe | Gly | Ser | Leu | Leu | His | Glu | | |
| | | 275 | | | | | 280 | | | | | 285 | | | | | |
| Phe | Gly | Leu | Leu | Glu | Ser | Pro | Val | Ala | Ala | Glu | Glu | Ala | Pro | Trp | Pro | | |
| | 290 | | | | | 295 | | | | | 300 | | | | | | |
| Pro | Pro | Glu | Gly | Ala | Phe | Val | Gly | Tyr | Val | Leu | Ser | Arg | Pro | Glu | Pro | | |
| 305 | | | | | 310 | | | | | 315 | | | | | 320 | | |
| Met | Trp | Ala | Glu | Leu | Asn | Ala | Leu | Ala | Ala | Ala | Trp | Gly | Gly | Arg | Val | | |
| | | | | 325 | | | | | 330 | | | | | 335 | | | |
| His | Arg | Ala | Ala | Asp | Pro | Leu | Ala | Gly | Leu | Lys | Asp | Leu | Lys | Glu | Val | | |
| | | | 340 | | | | | 345 | | | | | 350 | | | | |
| Arg | Gly | Leu | Leu | Ala | Lys | Asp | Leu | Ala | Val | Leu | Ala | Ser | Arg | Glu | Gly | | |
| | | 355 | | | | | 360 | | | | | 365 | | | | | |
| Leu | Asp | Leu | Val | Pro | Gly | Asp | Asp | Pro | Met | Leu | Leu | Ala | Tyr | Leu | Leu | | |
| | 370 | | | | | 375 | | | | | 380 | | | | | | |
| Asp | Pro | Ser | Asn | Thr | Thr | Pro | Glu | Gly | Val | Ala | Arg | Arg | Tyr | Gly | Gly | | |
| 385 | | | | | 390 | | | | | 395 | | | | | 400 | | |
| Glu | Trp | Thr | Glu | Asp | Ala | Ala | His | Arg | Ala | Leu | Leu | Ser | Glu | Arg | Leu | | |
| | | | | 405 | | | | | 410 | | | | | 415 | | | |
| His | Arg | Asn | Leu | Leu | Lys | Arg | Leu | Glu | Gly | Glu | Glu | Lys | Leu | Leu | Trp | | |
| | | | 420 | | | | | 425 | | | | | 430 | | | | |
| Leu | Tyr | His | Glu | Val | Glu | Lys | Pro | Leu | Ser | Arg | Val | Leu | Ala | His | Met | | |
| | | 435 | | | | | 440 | | | | | | 445 | | | | |
| Glu | Ala | Thr | Gly | Val | Arg | Arg | Asp | Val | Ala | Tyr | Leu | Gln | Ala | Leu | Ser | | |
| | 450 | | | | | 455 | | | | | 460 | | | | | | |
| Leu | Glu | Leu | Ala | Glu | Glu | Ile | Arg | Arg | Leu | Glu | Glu | Glu | Val | Phe | Arg | | |
| 465 | | | | | 470 | | | | | 475 | | | | | 480 | | |
| Leu | Ala | Gly | His | Pro | Phe | Asn | Leu | Asn | Ser | Arg | Asp | Gln | Leu | Glu | Arg | | |
| | | | | 485 | | | | | 490 | | | | | 495 | | | |
| Val | Leu | Phe | Asp | Glu | Leu | Arg | Leu | Pro | Ala | Leu | Lys | Lys | Thr | Lys | Lys | | |
| | | | 500 | | | | | 505 | | | | | 510 | | | | |
| Thr | Gly | Lys | Arg | Ser | Thr | Ser | Ala | Ala | Val | Leu | Glu | Ala | Leu | Arg | Glu | | |
| | | 515 | | | | | 520 | | | | | 525 | | | | | |
| Ala | His | Pro | Ile | Val | Glu | Lys | Ile | Leu | Gln | His | Arg | Glu | Leu | Thr | Lys | | |
| 535 | | | | | 540 | | | | | | | | | | | | 530 |
| Leu | Lys | Asn | Thr | Tyr | Val | Asp | Pro | Leu | Pro | Ser | Leu | Val | His | Pro | Arg | | |

<211> 27
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 473
 aagccactcc ggggtgatca ggtaacc 27

 <210> 474
 <211> 27
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 474
 atcaccccg agtggccttg ggagaag 27

 <210> 475
 <211> 2514
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 475
 atgaattccc tgcccctctt tgagcccaag ggccgggtgc ttctggtgga cggccaccac 60
 ctggcctacc gtaccttttt tgccctgaag ggccctacca ccagccgcgg ggagccggtc 120
 caggcgggtgt acgggttttg caagagcctt ttgaaggcgc taagggaaga cggggatgtg 180
 gtgatcgtgg tctttgacgc cgaggccccc tccttcgcc accagaccta cgaggcctac 240
 aaggcggggc gggctccac ccccgaggac tttcccggc agcttgccct tatcaaggag 300
 atggtggacc ttttgggcct ggagcgcctc gaggtgccgg gctttgaagc ggatgacgtc 360
 ctggctaccc tggccaagaa ggcggaagaa gaaggctacg aagtgcgcac cctcaccgcg 420
 gaccgggacc tttaccagct tctttcggag cgaatctcca tccttcaccc ggaggggttac 480
 ctgatcacc cggagtggct ttgggagaag tacggcctca ggccggagca gtgggtggac 540
 ttccgcgccc tcgtggggga cccctccgac aacctcccc ggggtcaagg catcggggag 600
 aagaccgccc tcaagctcct caaggagtgg ggaagcctgg aaaacctcct caagaacctg 660
 gaccgggtaa agccagaaaa cgtccgggag aagatcaagg cccacctgga agacctcagg 720
 ctctccttgg agctctcccg ggtgcgcacc gacctcccc tggaggtgga cctcgcccag 780

| | | | | | | |
|-------------|-------------|-------------|------------|------------|-------------|------|
| gggcgggagc | ccgaccggga | ggggcttagg | gccttcctgg | agaggctgga | gttcggcagc | 840 |
| ctcctccacg | agttcggcct | cctggaggcc | cccgcccccc | tggaggaggc | cccctggccc | 900 |
| ccgccggaag | gggccttcgt | gggcttcgtc | ctctcccgcc | ccgagcccat | gtgggaggag | 960 |
| cttaaagccc | tggccgcctg | caggggaggc | cgctgcacc | gggcagcaga | ccccttggcg | 1020 |
| gggctaaagg | acctcaagga | ggtccggggc | ctcctcgcca | aggacctcgc | cgtcttggcc | 1080 |
| tcgagggagg | ggctagacct | cgtgcccggg | gacgacccca | tgtcctcgc | ctacctcctg | 1140 |
| gacccttcga | acaccacccc | cgaggggggtg | gcgcggcgct | acggggggga | gtggacggag | 1200 |
| gacgccgccc | accgggcccct | cctctcgag | aggctccatc | ggaacctcct | taagcgccctc | 1260 |
| gagggggagg | agaagctcct | ttggctctac | cacgagggtg | aaaagcccct | ctcccggtc | 1320 |
| ctggcccata | tggaggccac | cgggggtacg | cgggacgtgg | cctaccttca | ggccctttcc | 1380 |
| ctggagcttg | cggaggagat | ccgccgcctc | gaggaggagg | tcttcgcctt | ggcgggccac | 1440 |
| cccttcaacc | tcaactcccg | ggaccagctg | gaaagggtgc | tctttgacga | gcttaggctt | 1500 |
| cccgcttga | agaagacgaa | gaagacaggc | aagcgctcca | ccagcgccgc | ggtgctggag | 1560 |
| gccctacggg | aggcccaccc | catcgtggag | aagatcctcc | agcaccggga | gctaccaag | 1620 |
| ctcaagaaca | cctacgtgga | ccccctcca | agcctcgtcc | accgaggac | gggccgcctc | 1680 |
| cacaccgct | tcaaccagac | ggccacggcc | acggggaggc | ttagtagctc | cgacccaac | 1740 |
| ctgcagaaca | tccccgtccg | caccccttg | ggccagagga | tccgcgggc | cttcgtggcc | 1800 |
| gaggcgggtt | gggcgttgg | ggccctggac | tatagccaga | tagagctccg | cgtcctcgcc | 1860 |
| cacctctccg | gggacgaaaa | cctgatcagg | gtcttccagg | aggggaagga | catccacacc | 1920 |
| cagaccgcaa | gctggatggt | cggcgtcccc | ccggaggccg | tggaccccct | gatgcgccgg | 1980 |
| gcggccaaga | cggatgaactt | cggcgtcctc | tacggcatgt | ccgcccatag | gctctcccag | 2040 |
| gagcttgcca | tccccctacga | ggaggcgggtg | gcctttatag | agcgctactt | ccaaagcttc | 2100 |
| cccaagggtgc | gggcctggat | agaaaagacc | ctggaggagg | ggaggaagcg | gggctacgtg | 2160 |
| gaaaccctct | tcggaagaag | gcgctacgtg | cccgacctca | acgcccgggt | gaagagcgtc | 2220 |
| agggaggccg | cggagcgcat | ggccttcaac | atgcccgtcc | agggcaccgc | cgccgacctc | 2280 |
| atgaagctcg | ccatggtgaa | gctcttcccc | cgctccggg | agatgggggc | cgcgatgctc | 2340 |
| ctccaggctcg | ccaacgagct | cctcctggag | gcccccaag | cgcgggccga | ggaggtggcg | 2400 |
| gctttggcca | aggaggccat | ggagaaggcc | tatccctcgc | ccgtgcccct | ggaggtggag | 2460 |

gtggggatgg gggaggactg gctttccgcc aagggtcacc accaccacca ccac

2514

<210> 476
<211> 838
<212> PRT
<213> Artificial

<220>
<223> Synthetic

<400> 476

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Asn | Ser | Leu | Pro | Leu | Phe | Glu | Pro | Lys | Gly | Arg | Val | Leu | Leu | Val | 1 | 5 | 10 | 15 |
| Asp | Gly | His | His | Leu | Ala | Tyr | Arg | Thr | Phe | Phe | Ala | Leu | Lys | Gly | Leu | 20 | 25 | 30 | |
| Thr | Thr | Ser | Arg | Gly | Glu | Pro | Val | Gln | Ala | Val | Tyr | Gly | Phe | Ala | Lys | 35 | 40 | 45 | |
| Ser | Leu | Leu | Lys | Ala | Leu | Arg | Glu | Asp | Gly | Asp | Val | Val | Ile | Val | Val | 50 | 55 | 60 | |
| Phe | Asp | Ala | Glu | Ala | Pro | Ser | Phe | Arg | His | Gln | Thr | Tyr | Glu | Ala | Tyr | 65 | 70 | 75 | 80 |
| Lys | Ala | Gly | Arg | Ala | Pro | Thr | Pro | Glu | Asp | Phe | Pro | Arg | Gln | Leu | Ala | 85 | 90 | 95 | |
| Leu | Ile | Lys | Glu | Met | Val | Asp | Leu | Leu | Gly | Leu | Glu | Arg | Leu | Glu | Val | 100 | 105 | 110 | |
| Pro | Gly | Phe | Glu | Ala | Asp | Asp | Val | Leu | Ala | Thr | Leu | Ala | Lys | Lys | Ala | 115 | 120 | 125 | |
| Glu | Lys | Glu | Gly | Tyr | Glu | Val | Arg | Ile | Leu | Thr | Ala | Asp | Arg | Asp | Leu | 130 | 135 | 140 | |
| Tyr | Gln | Leu | Leu | Ser | Glu | Arg | Ile | Ser | Ile | Leu | His | Pro | Glu | Gly | Tyr | 145 | 150 | 155 | 160 |
| Leu | Ile | Thr | Pro | Glu | Trp | Leu | Trp | Glu | Lys | Tyr | Gly | Leu | Arg | Pro | Glu | 165 | 170 | 175 | |
| Gln | Trp | Val | Asp | Phe | Arg | Ala | Leu | Val | Gly | Asp | Pro | Ser | Asp | Asn | Leu | 180 | 185 | 190 | |
| Pro | Gly | Val | Lys | Gly | Ile | Gly | Glu | Lys | Thr | Ala | Leu | Lys | Leu | Leu | Lys | 195 | 200 | 205 | |
| Glu | Trp | Gly | Ser | Leu | Glu | Asn | Leu | Leu | Lys | Asn | Leu | Asp | Arg | Val | Lys | 210 | 215 | 220 | |
| Pro | Glu | Asn | Val | Arg | Glu | Lys | Ile | Lys | Ala | His | Leu | Glu | Asp | Leu | Arg | 225 | 230 | 235 | 240 |

Tyr Val Asp Pro Leu Pro Ser Leu Val His Pro Arg Thr Gly Arg Leu
 545 550 555 560
 His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser Ser
 565 570 575
 Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly Gln
 580 585 590
 Arg Ile Arg Arg Ala Phe Val Ala Glu Ala Gly Trp Ala Leu Val Ala
 595 600 605
 Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser Gly
 610 615 620
 Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys Asp Ile His Thr
 625 630 635 640
 Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu Ala Val Asp Pro
 645 650 655
 Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly Val Leu Tyr Gly
 660 665 670
 Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu Glu
 675 680 685
 Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val Arg
 690 695 700
 Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys Arg Gly Tyr Val
 705 710 715 720
 Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn Ala Arg
 725 730 735
 Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro
 740 745 750
 Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu
 755 760 765
 Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu Leu Gln Val Ala
 770 775 780
 Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala Glu Glu Val Ala
 785 790 795 800
 Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro Leu Ala Val Pro
 805 810 815
 Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala Lys Gly
 820 825 830
 His His His His His
 835

<210> 477
 <211> 64
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 477
 caccaccacc accaccacgt cgactagtgc tagcgtcgac tagctgcagg catgcaagct 60
 tggc 64

 <210> 478
 <211> 64
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 478
 gccaaagcttg catgcctgca gctagtcgac gctagcacta gtcgacgtgg tgggtggtggt 60
 ggtg 64

 <210> 479
 <211> 41
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 479
 caggaagcgg ccgctcgac atgaccatga ttacgccaa g c 41

 <210> 480
 <211> 42
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 480
 gggcccgcca gggtcgactc agggcgatgg cccactacgt ga 42

 <210> 481
 <211> 3135
 <212> DNA
 <213> Artificial

<220>

<223> Synthetic

<400> 481

atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
gaaccggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac 180
gggtacaagg ccgtcttcgt ggtctttgac gccaaggccc cctccttcog ccacgaggcc 240
tacgaggcct acaaggcggg gagggccccg acccccaggg acttcccccg gcagctcgcc 300
ctcatcaagg agctggtgga cctcctgggg ttaccgcgc tcgagggtccc cggtacgag 360
gcggacgacg ttctcgccac cctggccaag aaggcggaag aggaggggta cgagggtgcgc 420
atcctcaccg ccgaccgga cctctaccaa ctgctctccg accgcgtcgc cgtcctccac 480
cccgagggcc acctcatcac ccggagtggt ctttgggaga agtacggcct caggccggag 540
cagtgggtgg acttccgcgc cctcgtgggg gacccctccg acaacctccc cggggtcaag 600
ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc 660
ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg 720
gaagacctca ggtctctctt ggagctctcc cgggtgcgca ccgacctccc cctggagggtg 780
gacctcgccc aggggcggga gcccgaccgg gaggggctta gggccttcct ggagaggctg 840
gagttcggca gcctcctcca cgagttcggc ctcttgaggg ccccgcccc cctggaggag 900
gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctcccg ccccgagccc 960
atgtgggcgg agcttaaagc cctggccgcc tgcaggggag gccgcgtgca ccgggcagca 1020
gaccccttgg cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc 1080
gccgtcttgg cctcgaggga ggggctagac ctggtgccc gggacgacct catgctctc 1140
gcctacctcc tggacccttc gaacaccacc cccgaggggg tggcgcggcg ctacgggggg 1200
gagtggacgg aggacgccgc ccaccgggcc ctctctcgg agaggctcca tcggaacctc 1260
cttaagcgcc tcgaggggga ggagaagctc ctttggctct accacgagggt ggaaaagccc 1320
ctctccccgg tcctggccca tatggaggcc accgggttac ggcgggacgt ggcctacctt 1380
caggcccttt ccttgagct tgccgaggag atccgccgcc tcgaggagga ggtcttccgc 1440
ttggcggggc accccttcaa cctcaactcc cgggaccagc tggaaagggt gctctttgac 1500
gagcttaggc ttccgcctt gaagaagacg aagaagacag gcaagcgctc caccagcgcc 1560

10434034350

<212> PRT
<213> Artificial

<220>
<223> Synthetic

<400> 482

Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
1 5 10 15
Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu
20 25 30
Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
35 40 45
Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala
50 55 60
Val Phe Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala
65 70 75 80
Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro
85 90 95
Arg Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr
100 105 110
Arg Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu
115 120 125
Ala Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala
130 135 140
Asp Arg Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His
145 150 155 160
Pro Glu Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly
165 170 175
Leu Arg Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro
180 185 190
Ser Asp Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu
195 200 205
Lys Leu Leu Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu
210 215 220
Asp Arg Val Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu
225 230 235 240
Glu Asp Leu Arg Leu Ser Leu Glu Leu Ser Arg Val Arg Thr Asp Leu
245 250 255
Pro Leu Glu Val Asp Leu Ala Gln Gly Arg Glu Pro Asp Arg Glu Gly


```

aacctgcaga acatccccgt ccgcaccccc ttgggccaga ggatccgccg ggccttcgtg 1800
gccgaggcgg gttgggcgtt ggtggccctg gactatagcc agatagagct ccgcgtcctc 1860
gccacctctt ccggggacga aaacctgatc agggctcttc aggaggggaa ggacatccac 1920
accagaccg caagctggat gttcggcgtc ccccgaggagg ccgtggaccc cctgatgcgc 1980
cgggcggcca agacggtgaa cttcggcgtc ctctacggca tgtccgcccc taggctctcc 2040
caggagcttg ccatccccta cgaggaggcg gtggccttta tagagcgcta cttccaaagc 2100
ttccccaagg tgcgggcctg gatagaaaag accctggagg aggggaggaa gcggggctac 2160
gtggaaaccc tcttcggaag aaggcgctac gtgcccgacc tcaacgcccg ggtgaagagc 2220
gtcagggagg ccgcggagcg catggccttc aacatgcccg tccagggcac cgcgcgcgac 2280
ctcatgaagc tcgccatggt gaagctcttc cccgcctcc gggagatggg ggcccgcgatg 2340
ctcctccagg tcgccaacga gctcctcctg gagggccccc aagcgcgggc cgaggagggtg 2400
gcggcttttg ccaaggaggc catggagaag gcctatcccc tcgccgtgcc cctggagggtg 2460
gaggtgggga tgggggagga ctggctttcc gccaaagggtc accaccacca ccaccac 2517

```

```

<210> 485
<211> 839
<212> PRT
<213> Artificial

```

```

<220>
<223> Synthetic

```

```

<400> 485

```

```

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1          5          10          15
Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
20        25        30
Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
35        40        45
Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
50        55        60
Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
65        70        75        80
Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
85        90        95
Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr Arg Leu
100       105       110

```

Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
 115 120 125
 Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
 130 135 140
 Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
 145 150 155 160
 Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
 165 170 175
 Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
 180 185 190
 Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
 195 200 205
 Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
 210 215 220
 Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
 225 230 235 240
 Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
 245 250 255
 Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
 260 265 270
 Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
 275 280 285
 Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu
 290 295 300
 Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala
 305 310 315 320
 Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala
 325 330 335
 Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val Arg Gly Leu
 340 345 350
 Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly Leu Asp Leu
 355 360 365
 Val Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser
 370 375 380
 Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr
 385 390 395 400
 Glu Asp Ala Ala His Arg Ala Leu Leu Ser Glu Arg Leu His Arg Asn
 405 410 415

Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn Ala
725 730 735

Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
740 745 750

Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
755 760 765

Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu Leu Gln Val
770 775 780

Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala Glu Glu Val
785 790 795 800

Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Gly His His His His His His
835

<210> 486
<211> 36
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 486
gtggaccttc tgggctttac ccgcctcgag gccccg

36

<210> 487
<211> 36
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 487
cggggcctcg aggcgggtaa agcccagaag gtccac

36

<210> 488
<211> 842
<212> PRT
<213> Artificial

<220>
<223> Synthetic

<400> 488

Met Asn Ser Thr Pro Leu Phe Asp Leu Glu Glu Pro Pro Lys Arg Val
1 5 10 15

Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Tyr Ala Leu
20 25 30

Ser Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Met Val Tyr Gly Phe
35 40 45

Ala Arg Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Gln Ala Val Val
50 55 60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu
65 70 75 80

Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
85 90 95

Leu Ala Leu Val Lys Arg Leu Val Asp Leu Leu Gly Phe Thr Arg Leu
100 105 110

Glu Ala Pro Gly Tyr Glu Ala Asp Asp Val Leu Gly Thr Leu Ala Lys
115 120 125

Lys Ala Glu Arg Glu Gly Met Glu Val Arg Ile Leu Thr Gly Asp Arg
130 135 140

Asp Phe Phe Gln Leu Leu Ser Glu Lys Val Ser Val Leu Leu Pro Asp
145 150 155 160

Gly Thr Leu Val Thr Pro Lys Asp Val Gln Glu Lys Tyr Gly Val Pro
165 170 175

Pro Glu Arg Trp Val Asp Phe Arg Ala Leu Thr Gly Asp Arg Ser Asp
180 185 190

Asn Ile Pro Gly Val Ala Gly Ile Gly Glu Lys Thr Ala Leu Arg Leu
195 200 205

Leu Ala Glu Trp Gly Ser Val Glu Asn Leu Leu Lys Asn Leu Asp Arg
210 215 220

Val Lys Pro Asp Ser Leu Arg Arg Lys Ile Glu Ala His Leu Glu Asp
225 230 235 240

Leu His Leu Ser Leu Asp Leu Ala Arg Ile Arg Thr Asp Leu Pro Leu
245 250 255

Glu Val Asp Phe Lys Ala Leu Arg Arg Arg Thr Pro Asp Leu Glu Gly
260 265 270

Leu Arg Ala Phe Leu Glu Glu Leu Glu Phe Gly Ser Leu Leu His Glu
275 280 285

Phe Gly Leu Leu Gly Gly Glu Lys Pro Arg Glu Glu Ala Pro Trp Pro

| | | |
|---|-----|-----|
| 595 | 600 | 605 |
| Ala Leu Val Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala | | |
| 610 | 615 | 620 |
| His Leu Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys | | |
| 625 | 630 | 635 |
| Asp Ile His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu | | |
| | 645 | 650 |
| Ala Val Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly | | |
| | 660 | 665 |
| Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile | | |
| | 675 | 680 |
| Pro Tyr Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe | | |
| | 690 | 695 |
| Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys | | |
| | 705 | 710 |
| Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp | | |
| | 725 | 730 |
| Leu Asn Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala | | |
| | 740 | 745 |
| Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala | | |
| | 755 | 760 |
| Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu | | |
| | 770 | 775 |
| Leu Gln Val Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala | | |
| | 785 | 790 |
| Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro | | |
| | 805 | 810 |
| Leu Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu | | |
| | 820 | 825 |
| Ser Ala Lys Gly His His His His His His | | |
| | 835 | 840 |
| <210> 489 | | |
| <211> 24 | | |
| <212> DNA | | |
| <213> Artificial | | |
| <220> | | |
| <223> Synthetic | | |
| <400> 489 | | |
| tttaccgccc tcgaggtgcc gggc | | |

<210> 490
 <211> 36
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 490
 cggcacctcg aggcgggtaa agcccaaaag gtccac

36

<210> 491
 <211> 838
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 491

| | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| Met | Asn | Ser | Leu | Pro | Leu | Phe | Glu | Pro | Lys | Gly | Arg | Val | Leu | Leu | Val | |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | | |
| Asp | Gly | His | His | Leu | Ala | Tyr | Arg | Thr | Phe | Phe | Ala | Leu | Lys | Gly | Leu | |
| | | | 20 | | | | | 25 | | | | | 30 | | | |
| Thr | Thr | Ser | Arg | Gly | Glu | Pro | Val | Gln | Ala | Val | Tyr | Gly | Phe | Ala | Lys | |
| | | 35 | | | | | 40 | | | | | 45 | | | | |
| Ser | Leu | Leu | Lys | Ala | Leu | Arg | Glu | Asp | Gly | Asp | Val | Val | Ile | Val | Val | |
| | 50 | | | | | 55 | | | | | 60 | | | | | |
| Phe | Asp | Ala | Lys | Ala | Pro | Ser | Phe | Arg | His | Gln | Thr | Tyr | Glu | Ala | Tyr | |
| 65 | | | | 70 | | | | | | 75 | | | | | 80 | |
| Lys | Ala | Gly | Arg | Ala | Pro | Thr | Pro | Glu | Asp | Phe | Pro | Arg | Gln | Leu | Ala | |
| | | | 85 | | | | | | 90 | | | | | 95 | | |
| Leu | Ile | Lys | Glu | Met | Val | Asp | Leu | Leu | Gly | Phe | Thr | Arg | Leu | Glu | Val | |
| | | 100 | | | | | | 105 | | | | | 110 | | | |
| Pro | Gly | Phe | Glu | Ala | Asp | Asp | Val | Leu | Ala | Thr | Leu | Ala | Lys | Lys | Ala | |
| | | 115 | | | | | 120 | | | | | | 125 | | | |
| Glu | Lys | Glu | Gly | Tyr | Glu | Val | Arg | Ile | Leu | Thr | Ala | Asp | Arg | Asp | Leu | |
| | 130 | | | | | 135 | | | | | 140 | | | | | |
| Tyr | Gln | Leu | Leu | Ser | Glu | Arg | Ile | Ser | Ile | Leu | His | Pro | Glu | Gly | Tyr | |
| 145 | | | | 150 | | | | | | 155 | | | | | 160 | |
| Leu | Ile | Thr | Pro | Glu | Trp | Leu | Trp | Glu | Lys | Tyr | Gly | Leu | Lys | Pro | Ser | |
| | | | | 165 | | | | | 170 | | | | | 175 | | |

Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala Glu Glu Val Ala
785 790 795 800

Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro Leu Ala Val Pro
805 810 815

Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala Lys Gly
820 825 830

His His His His His His
835

<210> 492
<211> 34
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 492
atcgtggtct ttgacgccga ggccccctcc ttcc 34

<210> 493
<211> 34
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 493
ggaaggaggg ggcctcggcg tcaaagacca cgat 34

<210> 494
<211> 842
<212> PRT
<213> Artificial

<220>
<223> Synthetic

<400> 494

Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
1 5 10 15

Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu
20 25 30

Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
35 40 45

Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala
50 55 60

| | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| Val | Phe | Val | Val | Phe | Asp | Ala | Glu | Ala | Pro | Ser | Phe | Arg | His | Glu | Ala | |
| 65 | | | | | 70 | | | | | 75 | | | | | 80 | |
| Tyr | Glu | Ala | Tyr | Lys | Ala | Gly | Arg | Ala | Pro | Thr | Pro | Glu | Asp | Phe | Pro | |
| | | | | 85 | | | | | 90 | | | | | 95 | | |
| Arg | Gln | Leu | Ala | Leu | Ile | Lys | Glu | Leu | Val | Asp | Leu | Leu | Gly | Phe | Thr | |
| | | | 100 | | | | | 105 | | | | | 110 | | | |
| Arg | Leu | Glu | Val | Pro | Gly | Tyr | Glu | Ala | Asp | Asp | Val | Leu | Ala | Thr | Leu | |
| | | 115 | | | | | 120 | | | | | 125 | | | | |
| Ala | Lys | Lys | Ala | Glu | Lys | Glu | Gly | Tyr | Glu | Val | Arg | Ile | Leu | Thr | Ala | |
| | 130 | | | | | 135 | | | | | 140 | | | | | |
| Asp | Arg | Asp | Leu | Tyr | Gln | Leu | Val | Ser | Asp | Arg | Val | Ala | Val | Leu | His | |
| 145 | | | | | 150 | | | | | 155 | | | | | 160 | |
| Pro | Glu | Gly | His | Leu | Ile | Thr | Pro | Glu | Trp | Leu | Trp | Glu | Lys | Tyr | Gly | |
| | | | | 165 | | | | | 170 | | | | | 175 | | |
| Leu | Arg | Pro | Glu | Gln | Trp | Val | Asp | Phe | Arg | Ala | Leu | Val | Gly | Asp | Pro | |
| | | | 180 | | | | | 185 | | | | | 190 | | | |
| Ser | Asp | Asn | Leu | Pro | Gly | Val | Lys | Gly | Ile | Gly | Glu | Lys | Thr | Ala | Leu | |
| | | 195 | | | | | 200 | | | | | 205 | | | | |
| Lys | Leu | Leu | Lys | Glu | Trp | Gly | Ser | Leu | Glu | Asn | Leu | Leu | Lys | Asn | Leu | |
| | 210 | | | | | 215 | | | | | 220 | | | | | |
| Asp | Arg | Val | Lys | Pro | Glu | Asn | Val | Arg | Glu | Lys | Ile | Lys | Ala | His | Leu | |
| 225 | | | | | 230 | | | | | 235 | | | | | 240 | |
| Glu | Asp | Leu | Arg | Leu | Ser | Leu | Glu | Leu | Ser | Arg | Val | Arg | Thr | Asp | Leu | |
| | | | | 245 | | | | | 250 | | | | | 255 | | |
| Pro | Leu | Glu | Val | Asp | Leu | Ala | Gln | Gly | Arg | Glu | Pro | Asp | Arg | Glu | Gly | |
| | | | 260 | | | | | 265 | | | | | 270 | | | |
| Leu | Arg | Ala | Phe | Leu | Glu | Arg | Leu | Glu | Phe | Gly | Ser | Leu | Leu | His | Glu | |
| | | | 275 | | | | 280 | | | | | 285 | | | | |
| Phe | Gly | Leu | Leu | Glu | Ala | Pro | Ala | Pro | Leu | Glu | Glu | Ala | Pro | Trp | Pro | |
| | 290 | | | | | 295 | | | | | 300 | | | | | |
| Pro | Pro | Glu | Gly | Ala | Phe | Val | Gly | Phe | Val | Leu | Ser | Arg | Pro | Glu | Pro | |
| 305 | | | | | 310 | | | | | 315 | | | | | 320 | |
| Met | Trp | Ala | Glu | Leu | Lys | Ala | Leu | Ala | Ala | Cys | Arg | Gly | Gly | Arg | Val | |
| | | | | 325 | | | | | 330 | | | | | 335 | | |
| His | Arg | Ala | Ala | Asp | Pro | Leu | Ala | Gly | Leu | Lys | Asp | Leu | Lys | Glu | Val | |
| | | | 340 | | | | | 345 | | | | | 350 | | | |
| Arg | Gly | Leu | Leu | Ala | Lys | Asp | Leu | Ala | Val | Leu | Ala | Ser | Arg | Glu | Gly | |
| | | 355 | | | | | 360 | | | | | 365 | | | | |

Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile
675 680 685

Pro Tyr Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe
690 695 700

Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys
705 710 715 720

Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp
725 730 735

Leu Asn Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala
740 745 750

Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala
755 760 765

Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu
770 775 780

Leu Gln Val Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala
785 790 795 800

Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro
805 810 815

Leu Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu
820 825 830

Ser Ala Lys Gly His His His His His His
835 840

<210> 495
<211> 2517
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 495
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag cgggggggag 120
ccggtgcagg cggctctacgg cttoGCCaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgcGag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggccgggc cccacgcGg gaggactttc cccggcaact cgcctcatc 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcacctc 420

| | |
|--|------|
| accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag | 480 |
| gggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg | 540 |
| gccgactacc gggccctgac cggggacgag tccgacaacc ttcccgggggt caagggcatc | 600 |
| ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag | 660 |
| aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccatat ggacgatctg | 720 |
| aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggagggt ggacttcgcc | 780 |
| aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc | 840 |
| agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga gggccctgg | 900 |
| cccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggccc | 960 |
| gatcttcttg ccctggccgc cgccaggggc ggccgcgtgc accgggcagc agaccccttg | 1020 |
| gcggggctaa aggacctcaa ggaggtccgg ggctcctcg ccaaggacct cgccgtcttg | 1080 |
| gcctcgaggg aggggctaga cctcgtgccc ggggacgacc ccatgctcct cgctacctc | 1140 |
| ctggaccctt cgaacaccac ccccgagggg gtggcgcggc gctacggggg ggagtggacg | 1200 |
| gaggacgccg cccaccgggc cctcctctcg gagaggctcc atcggaacct ccttaagcgc | 1260 |
| ctcgaggggg aggagaagct cctttggctc taccacgagg tggaaaagcc cctctcccgg | 1320 |
| gtcctggccc atatggaggc caccggggta cggcgggacg tggcctacct tcaggccctt | 1380 |
| tccctggagc ttgcggagga gatccgccgc ctcgaggagg aggtcttccg cttggcgggc | 1440 |
| caccccttca acctcaactc ccgggaccag ctggaaaggg tgctctttga cgagcttagg | 1500 |
| cttcccgcct tgaagaagac gaagaagaca ggcaagcgt ccaccagcgc cgcggtgctg | 1560 |
| gaggccctac gggaggccca ccccatcgtg gagaagatcc tccagcaccg ggagctcacc | 1620 |
| aagctcaaga acacctacgt ggacccccct ccaagcctcg tccacccgag gacgggcccgc | 1680 |
| ctccacaccc gcttcaacca gacggccacg gccacgggga ggcttagtag ctccgacccc | 1740 |
| aacctgcaga acatccccgt ccgcaccccc ttggggccaga ggatccgccg ggccttcgtg | 1800 |
| gccgaggcgg gttgggcgtt ggtggccctg gactatagcc agatagagct ccgcgtcctc | 1860 |
| gccacctct ccggggacga aaacctgatc agggctcttc aggaggggaa ggacatccac | 1920 |
| accagaccg caagctggat gttcggcgtc cccccggagg ccgtggaccc cctgatgcgc | 1980 |
| cgggcggcca agacggtgaa cttcggcgtc ctctacggca tgtccgcca taggctctcc | 2040 |
| caggagcttg ccatccccta cgaggaggcg gtggccttta tagagcgcta cttccaaagc | 2100 |
| ttccccaagg tgcgggcctg gatagaaaag accctggagg aggggaggaa gcggggctac | 2160 |

gtggaaaccc tcttcggaag aaggcgctac gtgcccgacc tcaacgcccg ggtgaagagc 2220
gtcagggagg ccgcgaggcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
ctcatgaagc tcgccatggt gaagctcttc ccccgccctcc gggagatggg ggcccgcgatg 2340
ctcctccagg tcgccaacga gctcctcctg gagggccccc aagcgcgggc cgaggagggtg 2400
gcggcctttgg ccaaggaggc catggagaag gcctatcccc tcgccgtgcc cctggagggtg 2460
gaggtggggga tgggggagga ctggccttcc gccaaagggtc accaccacca ccaccac 2517

<210> 496
<211> 839
<212> PRT
<213> Artificial

<220>
<223> Synthetic

<400> 496

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Asn | Ser | Gly | Met | Leu | Pro | Leu | Phe | Glu | Pro | Lys | Gly | Arg | Val | Leu |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Leu | Val | Asp | Gly | His | His | Leu | Ala | Tyr | Arg | Thr | Phe | His | Ala | Leu | Lys |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Gly | Leu | Thr | Thr | Ser | Arg | Gly | Glu | Pro | Val | Gln | Ala | Val | Tyr | Gly | Phe |
| | | | 35 | | | | 40 | | | | | 45 | | | |
| Ala | Lys | Ser | Leu | Leu | Lys | Ala | Leu | Lys | Glu | Asp | Gly | Asp | Ala | Val | Ile |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| Val | Val | Phe | Asp | Ala | Glu | Ala | Pro | Ser | Phe | Arg | His | Glu | Ala | Tyr | Gly |
| 65 | | | | | 70 | | | | 75 | | | | | 80 | |
| Gly | Tyr | Lys | Ala | Gly | Arg | Ala | Pro | Thr | Pro | Glu | Asp | Phe | Pro | Arg | Gln |
| | | | | 85 | | | | | 90 | | | | | 95 | |
| Leu | Ala | Leu | Ile | Lys | Glu | Leu | Val | Asp | Leu | Leu | Gly | Leu | Ala | Arg | Leu |
| | | | 100 | | | | | 105 | | | | | 110 | | |
| Glu | Val | Pro | Gly | Tyr | Glu | Ala | Asp | Asp | Val | Leu | Ala | Ser | Leu | Ala | Lys |
| | | 115 | | | | | 120 | | | | | 125 | | | |
| Lys | Ala | Glu | Lys | Glu | Gly | Tyr | Glu | Val | Arg | Ile | Leu | Thr | Ala | Asp | Lys |
| | 130 | | | | | 135 | | | | | 140 | | | | |
| Asp | Leu | Tyr | Gln | Leu | Leu | Ser | Asp | Arg | Ile | His | Val | Leu | His | Pro | Glu |
| 145 | | | | | 150 | | | | 155 | | | | | 160 | |
| Gly | Tyr | Leu | Ile | Thr | Pro | Ala | Trp | Leu | Trp | Glu | Lys | Tyr | Gly | Leu | Arg |
| | | | | 165 | | | | 170 | | | | | | 175 | |

| | | | |
|---|-----|-----|-----|
| Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp | 180 | 185 | 190 |
| Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu | 195 | 200 | 205 |
| Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg | 210 | 215 | 220 |
| Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu | 225 | 230 | 235 |
| Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu | 245 | 250 | 255 |
| Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala | 260 | 265 | 270 |
| Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu | 275 | 280 | 285 |
| Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu | 290 | 295 | 300 |
| Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala | 305 | 310 | 315 |
| Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala | 325 | 330 | 335 |
| Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val Arg Gly Leu | 340 | 345 | 350 |
| Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly Leu Asp Leu | 355 | 360 | 365 |
| Val Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser | 370 | 375 | 380 |
| Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr | 385 | 390 | 395 |
| Glu Asp Ala Ala His Arg Ala Leu Leu Ser Glu Arg Leu His Arg Asn | 405 | 410 | 415 |
| Leu Leu Lys Arg Leu Glu Gly Glu Glu Lys Leu Leu Trp Leu Tyr His | 420 | 425 | 430 |
| Glu Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met Glu Ala Thr | 435 | 440 | 445 |
| Gly Val Arg Arg Asp Val Ala Tyr Leu Gln Ala Leu Ser Leu Glu Leu | 450 | 455 | 460 |
| Ala Glu Glu Ile Arg Arg Leu Glu Glu Glu Val Phe Arg Leu Ala Gly | 465 | 470 | 475 |
| | | | 480 |

Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala Glu Glu Val
785 790 795 800

Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Gly His His His His His His
835

<210> 497
<211> 2526
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 497
atgaattcca cccactttt tgacctggag gaacccccca agcgggtgct tctgggtggac 60
ggccaccacc tggcctaccg caccttctat gccctgagcc tcaccacctc ccgggggggag 120
ccggtgcaga tgggtctacgg cttcgcccgg agcctcctca aggccttgaa ggaggacgga 180
caggcgggtgg tcgtggtctt tgacgccgag gccccctcct tccgccacga ggccctacgag 240
gcctacaagg cgggccgggc cccacccccg gaggacttcc cccgccagct cgccttggtc 300
aagcggctgg tggaccttct gggcctggtc cgcctcgagg ccccggggta cgaggcggac 360
gacgtcctgg gcacctggc caagaaggcc gaaagggagg ggatggaggt gcgcacctc 420
acgggagacc gggacttctt ccagctcctc tccgagaagg tctcggtcct cctgccggac 480
gggaccctgg tcaccccaa ggacgtccag gagaagtacg gggtgcccc ggagcgctgg 540
gtggacttcc gcgcctcac gggggaccgc tcggacaaca tccccggggt ggcggggata 600
ggggagaaga ccgcccttcg actcctcgca gagtggggga gcgtggaaaa cctcctgaag 660
aacctggacc gggtaaagcc ggactcgctc cggcgcaaga tagaggcgca cctcgaggac 720
ctccacctct ccttagacct ggcccgcatc cgcaccgacc tccccctgga ggtggacttt 780
aagggcctgc gccgcaggac ccccgacctg gagggcctga gggccttttt ggaggagctg 840
gagttcgga gacctctcca cgagttcggc ctctggggag gggagaagcc ccgggaggag 900
gccccctggc ccccgcccga aggggccttc gtgggcttcc tcctttcccg caaggagccc 960
atgtgggcgg agcttctggc cctggcgggc gcctcggggc gccgcgtgca ccgggcagca 1020
gacctcttgg cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc 1080

| | |
|--|------|
| gccgtcttgg cctcgaggga ggggctagac ctctgccccg gggacgaccc catgctcctc | 1140 |
| gcctacctcc tggacccttc gaacaccacc cccgaggggg tggcgcggcg ctacgggggg | 1200 |
| gagtggacgg aggacgccgc ccaccgggcc ctctctctcg agaggctcca tcggaacctc | 1260 |
| cttaagcgcc tcgaggggga ggagaagctc ctttggtctc accacgaggt ggaaaagccc | 1320 |
| ctctccccgg tcttgggcca tatggaggcc accggggtac ggcgggacgt ggcctacctt | 1380 |
| caggcccttt ccctggagct tgcggaggag atccgccgcc tcgaggagga ggtcttcgcg | 1440 |
| ttggcgggcc accccttcaa cctcaactcc cgggaccagc tggaaagggg gctctttgac | 1500 |
| gagcttaggc ttccgcctt gaagaagacg aagaagacag gcaagcgctc caccagcgcc | 1560 |
| gcggtgctgg aggccctacg ggaggccac cccatcgtgg agaagatcct ccagcaccgg | 1620 |
| gagctacca agctcaagaa cacctacgtg gacccctcc caagcctcgt ccacccgagg | 1680 |
| acgggcccgc tccacaccg cttcaaccag acggccacgg ccacggggag gcttagtagc | 1740 |
| tccgaccca acctgcagaa catccccgtc cgcacccct tgggccagag gatccgccgg | 1800 |
| gccttcgtgg ccgaggcggg ttgggcgttg gtggcccttg actatagcca gatagagctc | 1860 |
| cgcgtcctcg ccacctctc cggggacgaa aacctgatca ggtcttcca ggaggggaag | 1920 |
| gacatccaca ccagaccgc aagctggatg ttcggcgctc ccccgaggc cgtggacccc | 1980 |
| ctgatgcgcc gggcgccaa gacggtgaac ttcggcgctc tctacggcat gtccgccc | 2040 |
| aggctctccc aggagcttgc catcccctac gaggaggcgg tggcctttat agagcgctac | 2100 |
| ttccaaagct tccccaggt gcgggccttg atagaaaaga ccctggagga ggggaggaag | 2160 |
| cggggctacg tggaaaccct cttcggaaga aggcgctac tgcccacact caacgcccgg | 2220 |
| gtgaagagcg tcagggaggc cgcggagcgc atggccttca acatgcccg ccagggcacc | 2280 |
| gccgcccacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg | 2340 |
| gcccgcacgc tctccaggt cgccaacgag ctctctcttg agggccccca agcgcgggcc | 2400 |
| gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatcccct cgccgtgccc | 2460 |
| ctggaggtgg aggtggggat gggggaggac tggttttccg ccaaggggtca ccaccaccac | 2520 |
| caccac | 2526 |

<210> 498
 <211> 842
 <212> PRT
 <213> Artificial

 <220>

<223> Synthetic

<400> 498

Met Asn Ser Thr Pro Leu Phe Asp Leu Glu Glu Pro Pro Lys Arg Val
1 5 10 15

Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Tyr Ala Leu
20 25 30

Ser Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Met Val Tyr Gly Phe
35 40 45

Ala Arg Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Gln Ala Val Val
50 55 60

Val Val Phe Asp Ala Glu Ala Pro Ser Phe Arg His Glu Ala Tyr Glu
65 70 75 80

Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
85 90 95

Leu Ala Leu Val Lys Arg Leu Val Asp Leu Leu Gly Leu Val Arg Leu
100 105 110

Glu Ala Pro Gly Tyr Glu Ala Asp Asp Val Leu Gly Thr Leu Ala Lys
115 120 125

Lys Ala Glu Arg Glu Gly Met Glu Val Arg Ile Leu Thr Gly Asp Arg
130 135 140

Asp Phe Phe Gln Leu Leu Ser Glu Lys Val Ser Val Leu Leu Pro Asp
145 150 155 160

Gly Thr Leu Val Thr Pro Lys Asp Val Gln Glu Lys Tyr Gly Val Pro
165 170 175

Pro Glu Arg Trp Val Asp Phe Arg Ala Leu Thr Gly Asp Arg Ser Asp
180 185 190

Asn Ile Pro Gly Val Ala Gly Ile Gly Glu Lys Thr Ala Leu Arg Leu
195 200 205

Leu Ala Glu Trp Gly Ser Val Glu Asn Leu Leu Lys Asn Leu Asp Arg
210 215 220

Val Lys Pro Asp Ser Leu Arg Arg Lys Ile Glu Ala His Leu Glu Asp
225 230 235 240

Leu His Leu Ser Leu Asp Leu Ala Arg Ile Arg Thr Asp Leu Pro Leu
245 250 255

Glu Val Asp Phe Lys Ala Leu Arg Arg Arg Thr Pro Asp Leu Glu Gly
260 265 270

Leu Arg Ala Phe Leu Glu Glu Leu Glu Phe Gly Ser Leu Leu His Glu
275 280 285

<400> 499
atgaattccc tgcccctctt tgagcccaag ggccgggtgc ttctggtgga cggccaccac 60
ctggcctacc gtaccttttt tgccctgaag ggcctcacca ccagccgcgg ggagccggtc 120
caggcgggtgt acgggttttg caagagcctt ttgaaggcgc taagggaaga cggggatgtg 180
gtgatcgtgg tctttgacgc cgaggccccc tccttcgcgc accagaccta cgaggcctac 240
aaggcggggc gggctccac ccccgaggac tttcccggc agcttgccct tatcaaggag 300
atggtggacc ttttgggcct ggagcgctc gaggtgccgg gctttgaagc ggatgacgtc 360
ctggctaccc tggccaagaa ggcggaagaa gaaggtacg aagtgcgcac cctcaccgcg 420
gaccgggacc ttaccagct tctttcgag cgaatctcca tccttcaccc ggaggggttac 480
ctgatcacc cggagtggct ttgggagaag tatgggctta agccttccca gtgggtggac 540
taccgggcct tggccgggga cccttcgcac aacatccccg gcgtgaaggg catcggggag 600
aagacggcgg ccaagctgat ccgggagtgg ggaagcctgg aaaaccttct taagcacctg 660
gaacaggtga aacctgcctc cgtgcgggag aagatcctta gccacatgga ggacctcaag 720
ctatccctgg agctatccc ggtgcacacg gacttgctcc ttcaggtgga cttcgcccgg 780
cgccggggagc cggaccggga ggggcttaag gccttttttg agaggctgga gttcggaagc 840
ctcctccacg agttcggcct gttggaaagc ccggtggcgg cggaggaagc tccctggccg 900
cccccgagg gagccttcgt ggggtacgtt ctttccgcgc ccgagcccat gtgggcggag 960
cttaacgcct tggccgcgc ctggggcggc cgcgtgcacc gggcagcaga ccccttggcg 1020
gggctaaaag acctcaagga ggtccggggc ctctcgcca aggacctgc cgtcttggcc 1080
tcgagggagg ggctagacct cgtgcccggg gacgaccca tgctcctgc ctacctctg 1140
gaccttcga acaccacccc cgaggggggtg gcgcggcgct acggggggga gtggacggag 1200
gacgccgcc accgggccct cctctcgag aggtccatc ggaacctct taagcgctc 1260
gagggggagg agaagctcct ttggctctac cagaggtgg aaaagcccct ctcccgggtc 1320
ctggcccata tggaggccac cggggtacgg cgggacgtgg cctaccttca ggccctttcc 1380
ctggagcttg cggaggagat ccgccgcctc gaggaggagg tcttcgcgtt ggccggccac 1440
cccttcaacc tcaactccc ggaccagctg gaaaggggtg tctttgacga gcttaggctt 1500
ccgccttga agaagacgaa gaagacaggc aagcgctcca ccagcgccgc ggtgctggag 1560
gccctacggg aggccacccc catcgtggag aagatcctcc agcaccggga gctcaccaag 1620
ctcaagaaca cctacgtgga cccctccca agcctcgtcc acccgaggac gggccgcctc 1680

Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys Arg Gly Tyr Val
705 710 715 720

Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn Ala Arg
725 730 735

Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro
740 745 750

Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu
755 760 765

Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu Leu Gln Val Ala
770 775 780

Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala Glu Glu Val Ala
785 790 795 800

Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro Leu Ala Val Pro
805 810 815

Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala Lys Gly
820 825 830

His His His His His His
835

<210> 501
<211> 2517
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 501
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
ccggtgcagg cggtctacgg cttagccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccgag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggccgggc cccacgcgcg gaggactttc cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggttcacg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccag 480
gggtacctca tcaccccggc ctggctttgg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660

ctcctccagg tcgccaacga gctcctcctg gaggccccc aagcgcgggc cgaggaggtg 2400
 gcggcttttg ccaaggaggc catggagaag gcctatcccc tcgccgtgcc cctggaggtg 2460
 gaggtgggga tgggggagga ctggctttcc gccaaagggtc accaccacca ccaccac 2517

<210> 502
 <211> 839
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 502

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Asn | Ser | Gly | Met | Leu | Pro | Leu | Phe | Glu | Pro | Lys | Gly | Arg | Val | Leu |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Leu | Val | Asp | Gly | His | His | Leu | Ala | Tyr | Arg | Thr | Phe | His | Ala | Leu | Lys |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Gly | Leu | Thr | Thr | Ser | Arg | Gly | Glu | Pro | Val | Gln | Ala | Val | Tyr | Gly | Phe |
| | | | 35 | | | | 40 | | | | | 45 | | | |
| Ala | Lys | Ser | Leu | Leu | Lys | Ala | Leu | Lys | Glu | Asp | Gly | Asp | Ala | Val | Ile |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| Val | Val | Phe | Asp | Ala | Glu | Ala | Pro | Ser | Phe | Arg | His | Glu | Ala | Tyr | Gly |
| 65 | | | | | 70 | | | | 75 | | | | | 80 | |
| Gly | Tyr | Lys | Ala | Gly | Arg | Ala | Pro | Thr | Pro | Glu | Asp | Phe | Pro | Arg | Gln |
| | | | | 85 | | | | 90 | | | | | | 95 | |
| Leu | Ala | Leu | Ile | Lys | Glu | Leu | Val | Asp | Leu | Leu | Gly | Phe | Thr | Arg | Leu |
| | | | 100 | | | | | 105 | | | | | 110 | | |
| Glu | Val | Pro | Gly | Tyr | Glu | Ala | Asp | Asp | Val | Leu | Ala | Ser | Leu | Ala | Lys |
| | | | 115 | | | | 120 | | | | | 125 | | | |
| Lys | Ala | Glu | Lys | Glu | Gly | Tyr | Glu | Val | Arg | Ile | Leu | Thr | Ala | Asp | Lys |
| | | | 130 | | | | 135 | | | | 140 | | | | |
| Asp | Leu | Tyr | Gln | Leu | Leu | Ser | Asp | Arg | Ile | His | Val | Leu | His | Pro | Glu |
| 145 | | | | | 150 | | | | 155 | | | | | 160 | |
| Gly | Tyr | Leu | Ile | Thr | Pro | Ala | Trp | Leu | Trp | Glu | Lys | Tyr | Gly | Leu | Arg |
| | | | | 165 | | | | 170 | | | | | | 175 | |
| Pro | Asp | Gln | Trp | Ala | Asp | Tyr | Arg | Ala | Leu | Thr | Gly | Asp | Glu | Ser | Asp |
| | | | 180 | | | | | 185 | | | | | 190 | | |
| Asn | Leu | Pro | Gly | Val | Lys | Gly | Ile | Gly | Glu | Lys | Thr | Ala | Arg | Lys | Leu |
| | | | 195 | | | | 200 | | | | | 205 | | | |
| Leu | Glu | Glu | Trp | Gly | Ser | Leu | Glu | Ala | Leu | Leu | Lys | Asn | Leu | Asp | Arg |

| | | |
|---|-----|-------------|
| 515 | 520 | 525 |
| Ile Val Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys Leu Lys Asn | | |
| 530 | 535 | 540 |
| Thr Tyr Val Asp Pro Leu Pro Ser Leu Val His Pro Arg Thr Gly Arg | | |
| 545 | 550 | 555 560 |
| Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser | | |
| | 565 | 570 575 |
| Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly | | |
| | 580 | 585 590 |
| Gln Arg Ile Arg Arg Ala Phe Val Ala Glu Ala Gly Trp Ala Leu Val | | |
| | 595 | 600 605 |
| Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser | | |
| | 610 | 615 620 |
| Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys Asp Ile His | | |
| | 625 | 630 635 640 |
| Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu Ala Val Asp | | |
| | 645 | 650 655 |
| Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly Val Leu Tyr | | |
| | 660 | 665 670 |
| Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu | | |
| | 675 | 680 685 |
| Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val | | |
| | 690 | 695 700 |
| Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys Arg Gly Tyr | | |
| | 705 | 710 715 720 |
| Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn Ala | | |
| | 725 | 730 735 |
| Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met | | |
| | 740 | 745 750 |
| Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys | | |
| | 755 | 760 765 |
| Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu Leu Gln Val | | |
| | 770 | 775 780 |
| Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala Glu Glu Val | | |
| | 785 | 790 795 800 |
| Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro Leu Ala Val | | |
| | 805 | 810 815 |
| Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala Lys | | |

820

825

830

Gly His His His His His His
835

<210> 503

<211> 34

<212> DNA

<213> Artificial

<220>

<223> Synthetic

i

<400> 503

atcgtggtct ttgacgccga ggccccctcc ttcc

34

<210> 504

<211> 34

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 504

ggaaggagg ggccctcggcg tcaaagacca cgat

34

<210> 505

<211> 2520

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 505

atgaattcgg aggcgatgct gcccctcttt gagcccaagg gccgggtcct cctggtggac 60

ggccaccacc tggcctaccg caccttccac gccctgaagg gcctcaccac cagccggggg 120

gagccggtgc aggcggtcta cggttcgcc aagagcctcc tcaaggcct caaggaggac 180

ggggacgcgg tgatcgtggt ctttgacgcc gaggccccct cttccgccca cgaggcctac 240

ggggggtaca aggcggggccg ggccccccacg ccggaggact ttccccggca actcgcctc 300

atcaaggagc tgggtggacct cctgggggttc acgcgcctcg aggtcccggg ctacgaggcg 360

gacgacgtcc tggccagcct ggccaagaag gcggaaaagg agggctacga ggtccgcac 420

ctcaccgccg acaaagacct ttaccagctc ctttccgacc gcatccacgt cctccacccc 480

gaggggtacc tcatcacccc ggccctggctt tgggaaaagt acggcctgag gcccgaccag 540

tgggccgact accggggcct gaccggggac gaggccgaca accttcccgg ggtcaagggc 600

| | |
|---|------|
| atcggggaga agacggcgag gaagcttctg gaggagtggg ggagcctgga agccctcctc | 660 |
| aagaacctgg accggtgaa gcccgccatc cgggagaaga tcttggccca catggacgat | 720 |
| ctgaagctct cctgggacct ggccaaggtg cgcaccgacc tgcccttgga ggtggacttc | 780 |
| gccaaaaggc gggagccga cgggagagg cttaggccct ttctggagag gcttgagttt | 840 |
| ggcagcctcc tccacgagtt cggccttctg gaaagcccca aggccttgga ggaggcccc | 900 |
| tggccccgc cggaaggggc cttcgtgggc tttgtgcttt cccgcaagga gcccatgtgg | 960 |
| gccgatcttc tggccctggc cgcgcagcagg ggcgccgcg tgcaccgggc agcagacccc | 1020 |
| ttggcggggc taaaggacct caaggaggtc cggggcctcc tcgccaagga cctcgccgtc | 1080 |
| ttggcctcga gggaggggct agacctcgtg cccggggacg acccatgct cctcgccctac | 1140 |
| ctcctggacc cttcgaacac ccccccgag ggggtggcgc ggcgctacgg gggggagtgg | 1200 |
| acggaggacg ccgcccaccg ggccctcctc tcggagaggc tccatcgga cctccttaag | 1260 |
| cgctcagagg gggaggagaa gctcctttgg ctctaccacg aggtggaaaa gccctctcc | 1320 |
| cgggtcctgg cccatatgga ggccaccggg gtacggcggg acgtggccta ccttcaggcc | 1380 |
| ctttccctgg agcttgcgga ggagatccgc cgctcagagg aggaggtctt ccgcttggcg | 1440 |
| ggccaccct tcaacctcaa ctcccgggac cagctggaaa gggtgctctt tgacgagctt | 1500 |
| aggcttcccg ccttgaagaa gacgaagaag acaggcaagc gctccaccag cgcgcgggtg | 1560 |
| ctggaggccc tacgggaggc ccacccatc gtggagaaga tctccagca cggggagctc | 1620 |
| accaagctca agaacaccta cgtggacccc ctccaagcc tcgtccacc gaggacgggc | 1680 |
| cgctccaca cccgcttcaa ccagacggcc acggccacgg ggaggcttag tagctccgac | 1740 |
| cccaacctgc agaacatccc cgtccgcacc cccttgggcc agaggatccg cggggccttc | 1800 |
| gtggccgagg cgggttgggc gttggtggcc ctggactata gccagataga gctccgcgtc | 1860 |
| ctcgcaccac tctccgggga cgaaaacctg atcagggtct tccaggagg gaaggacatc | 1920 |
| cacaccaga ccgcaagctg gatgttcggc gtccccccgg aggcggtgga cccctgatg | 1980 |
| cgcggggcgg ccaagacggt gaacttcggc gtctctacg gcatgtccgc ccataggctc | 2040 |
| tcccaggagc ttgccatccc ctacgaggag gcggtggcct ttatagagcg ctacttccaa | 2100 |
| agcttcccca aggtgcgggc ctggatagaa aagacctgg aggaggggag gaagcggggc | 2160 |
| tacgtgaaaa ccctcttcgg aagaaggcgc tacgtgccc acctcaacgc cggggtgaag | 2220 |
| agcgtcaggg aggccgcgga gcgcatggcc ttcaacatgc ccgtccaggg caccgccgcc | 2280 |

gacctcatga agctcgccat ggtgaagctc ttcccccgcc tccgggagat gggggcccgc 2340
 atgctcctcc aggtcgccaa cgagctctc ctggaggccc cccaagcgcg ggccgaggag 2400
 gtggcggtt tggccaagga ggccatggag aaggcctatc ccctcgccgt gcccctggag 2460
 gtggaggtgg ggatggggga ggactggctt tccgccaagg gtcaccacca ccaccaccac 2520

<210> 506
 <211> 840
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 506

Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
 1 5 10 15
 Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu
 20 25 30
 Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
 35 40 45
 Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val
 50 55 60
 Ile Val Val Phe Asp Ala Glu Ala Pro Ser Phe Arg His Glu Ala Tyr
 65 70 75 80
 Gly Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg
 85 90 95
 Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr Arg
 100 105 110
 Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala
 115 120 125
 Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp
 130 135 140
 Lys Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro
 145 150 155 160
 Glu Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu
 165 170 175
 Arg Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser
 180 185 190
 Asp Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys
 195 200 205

Lys Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His
 515 520 525
 Pro Ile Val Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys Leu Lys
 530 535 540
 Asn Thr Tyr Val Asp Pro Leu Pro Ser Leu Val His Pro Arg Thr Gly
 545 550 555 560
 Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu
 565 570 575
 Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu
 580 585 590
 Gly Gln Arg Ile Arg Arg Ala Phe Val Ala Glu Ala Gly Trp Ala Leu
 595 600 605
 Val Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu
 610 615 620
 Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys Asp Ile
 625 630 635 640
 His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu Ala Val
 645 650 655
 Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly Val Leu
 660 665 670
 Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr
 675 680 685
 Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys
 690 695 700
 Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys Arg Gly
 705 710 715 720
 Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn
 725 730 735
 Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn
 740 745 750
 Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val
 755 760 765
 Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu Leu Gln
 770 775 780
 Val Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala Glu Glu
 785 790 795 800
 Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro Leu Ala
 805 810 815

Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala
820 825 830

Lys Gly His His His His His His
835 840

<210> 507
<211> 36
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 507
cagaccatga attcggaggc gatgctgccc ctcttt 36

<210> 508
<211> 36
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 508
aaagaggggc agcatcgcct ccgaattcat ggtctg 36

<210> 509
<211> 2517
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 509
atgaattcgg ggatgctgcc cctctttgag cccaagggcc ggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttctttgcc ctgaagggcc tcaccaccag ccgggggggag 120
ccggtgcagg cgggtctacgg ctctcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccgag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggccgggc ccccacgccg gaggactttc cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggttcacg cgctcagagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540

| | | | | | | |
|------------|-------------|-------------|------------|-------------|-------------|------|
| gccgactacc | gggccctgac | cggggacgag | tccgacaacc | ttcccgggggt | caagggcatc | 600 |
| ggggagaaga | cggcgaggaa | gcttctggag | gagtggggga | gcctggaagc | cctcctcaag | 660 |
| aacctggacc | ggctgaagcc | cgccatccgg | gagaagatcc | tggcccacat | ggacgatctg | 720 |
| aagctctcct | gggacctggc | caaggtgcgc | accgacctgc | ccctggagggt | ggacttcgcc | 780 |
| aaaaggcggg | agccccgaccg | ggagaggctt | agggcctttc | tggagagggt | tgagtttggc | 840 |
| agcctcctcc | acgagttcgg | ccttctggaa | agccccaagg | ccctggagga | ggccccctgg | 900 |
| cccccgccgg | aaggggcctt | cgtgggcctt | gtgctttccc | gcaaggagcc | catgtggggc | 960 |
| gatcttctgg | ccctggccgc | cgccaggggc | ggccgcgtgc | accgggcagc | agacccttg | 1020 |
| gcggggctaa | aggacctcaa | ggaggtccgg | ggcctcctcg | ccaaggacct | cgccgtcttg | 1080 |
| gcctcgaggg | aggggctaga | cctcgtgccc | ggggacgacc | ccatgctcct | cgcctacctc | 1140 |
| ctggaccctt | cgaacaccac | ccccgagggg | gtggcgcggc | gctacggggg | ggagtggacg | 1200 |
| gaggacgccg | cccaccgggc | cctcctctcg | gagaggctcc | atcggaacct | ccttaagcgc | 1260 |
| ctcgaggggg | aggagaagct | cctttggctc | taccacgagg | tggaaaagcc | cctctcccgg | 1320 |
| gtcctggccc | atatggaggc | caccggggta | cggcgggacg | tggcctacct | tcaggccctt | 1380 |
| tccttgagc | ttgcggagga | gatccgccgc | ctcgaggagg | aggtcttccg | cttggcgggc | 1440 |
| cacccttca | acctcaactc | ccgggaccag | ctggaaaggg | tgctctttga | cgagcttagg | 1500 |
| cttccgcct | tgaagaagac | gaagaagaca | ggcaagcgct | ccaccagcgc | cgcggtgctg | 1560 |
| gaggccctac | gggaggccca | ccccatcgtg | gagaagatcc | tccagcaccg | ggagctcacc | 1620 |
| aagctcaaga | acacctacgt | ggacccccctc | ccaagcctcg | tccacccgag | gacgggcccgc | 1680 |
| ctccacacc | gcttcaacca | gacggccacg | gccacgggga | ggcttagtag | ctccgacccc | 1740 |
| aacctgcaga | acatccccgt | ccgcaccccc | ttgggccaga | ggatccgccg | ggccttcgtg | 1800 |
| gccgaggcgg | gttgggcgtt | ggtggccctg | gactatagcc | agatagagct | ccgcgtcctc | 1860 |
| gccacctct | ccggggacga | aaacctgac | agggctcttc | aggaggggaa | ggacatccac | 1920 |
| accagaccg | caagctggat | gttcggcgtc | ccccggagg | ccgtggaccc | cctgatgcgc | 1980 |
| cgggcgggca | agacggtgaa | cttcggcgtc | ctctacggca | tgtccgcccc | taggctctcc | 2040 |
| caggagcttg | ccatccccct | cgaggaggcg | gtggccttta | tagagcgcta | cttccaaagc | 2100 |
| ttccccaagg | tgcgggcctg | gatagaaaag | accctggagg | aggggaggaa | gcggggctac | 2160 |
| gtggaaaccc | tcttcggaag | aaggcgctac | gtgcccgacc | tcaacgcccg | ggtgaagagc | 2220 |

gtcagggagg ccgcggagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
ctcatgaagc tcgccatggt gaagctcttc cccgcctcc gggagatggg ggcccgcgatg 2340
ctcctccagg tcgccaacga gctcctcctg gagggccccc aagcgcgggc cgaggaggtg 2400
gcggttttg ccaaggaggc catggagaag gcctatcccc tcgccgtgcc cctggaggtg 2460
gaggtgggga tgggggagga ctggctttcc gccaaagggtc accaccacca ccaccac 2517

<210> 510
<211> 839
<212> PRT
<213> Artificial

<220>
<223> Synthetic

<400> 510

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Asn | Ser | Gly | Met | Leu | Pro | Leu | Phe | Glu | Pro | Lys | Gly | Arg | Val | Leu |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Leu | Val | Asp | Gly | His | His | Leu | Ala | Tyr | Arg | Thr | Phe | Phe | Ala | Leu | Lys |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Gly | Leu | Thr | Thr | Ser | Arg | Gly | Glu | Pro | Val | Gln | Ala | Val | Tyr | Gly | Phe |
| | | | 35 | | | | 40 | | | | | 45 | | | |
| Ala | Lys | Ser | Leu | Leu | Lys | Ala | Leu | Lys | Glu | Asp | Gly | Asp | Ala | Val | Ile |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| Val | Val | Phe | Asp | Ala | Glu | Ala | Pro | Ser | Phe | Arg | His | Glu | Ala | Tyr | Gly |
| 65 | | | | | 70 | | | | 75 | | | | | 80 | |
| Gly | Tyr | Lys | Ala | Gly | Arg | Ala | Pro | Thr | Pro | Glu | Asp | Phe | Pro | Arg | Gln |
| | | | | 85 | | | | 90 | | | | | | 95 | |
| Leu | Ala | Leu | Ile | Lys | Glu | Leu | Val | Asp | Leu | Leu | Gly | Phe | Thr | Arg | Leu |
| | | | 100 | | | | | 105 | | | | | 110 | | |
| Glu | Val | Pro | Gly | Tyr | Glu | Ala | Asp | Asp | Val | Leu | Ala | Ser | Leu | Ala | Lys |
| | | 115 | | | | | 120 | | | | | 125 | | | |
| Lys | Ala | Glu | Lys | Glu | Gly | Tyr | Glu | Val | Arg | Ile | Leu | Thr | Ala | Asp | Lys |
| | | 130 | | | | 135 | | | | | 140 | | | | |
| Asp | Leu | Tyr | Gln | Leu | Leu | Ser | Asp | Arg | Ile | His | Val | Leu | His | Pro | Glu |
| 145 | | | | | 150 | | | | 155 | | | | | | 160 |
| Gly | Tyr | Leu | Ile | Thr | Pro | Ala | Trp | Leu | Trp | Glu | Lys | Tyr | Gly | Leu | Arg |
| | | | | 165 | | | | 170 | | | | | | 175 | |
| Pro | Asp | Gln | Trp | Ala | Asp | Tyr | Arg | Ala | Leu | Thr | Gly | Asp | Glu | Ser | Asp |
| | | | 180 | | | | | 185 | | | | | 190 | | |

| | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|--|
| Asp | Glu | Leu | Arg | Leu | Pro | Ala | Leu | Lys | Lys | Thr | Lys | Lys | Thr | Gly | Lys | | |
| | | | | 500 | | | | | | 505 | | | | | 510 | | |
| Arg | Ser | Thr | Ser | Ala | Ala | Val | Leu | Glu | Ala | Leu | Arg | Glu | Ala | His | Pro | | |
| | | 515 | | | | | | | | 520 | | | | 525 | | | |
| Ile | Val | Glu | Lys | Ile | Leu | Gln | His | Arg | Glu | Leu | Thr | Lys | Leu | Lys | Asn | | |
| | | 530 | | | | | 535 | | | | | 540 | | | | | |
| Thr | Tyr | Val | Asp | Pro | Leu | Pro | Ser | Leu | Val | His | Pro | Arg | Thr | Gly | Arg | | |
| 545 | | | | | | 550 | | | | | 555 | | | | 560 | | |
| Leu | His | Thr | Arg | Phe | Asn | Gln | Thr | Ala | Thr | Ala | Thr | Gly | Arg | Leu | Ser | | |
| | | | | 565 | | | | | 570 | | | | | | 575 | | |
| Ser | Ser | Asp | Pro | Asn | Leu | Gln | Asn | Ile | Pro | Val | Arg | Thr | Pro | Leu | Gly | | |
| | | | 580 | | | | | 585 | | | | | | 590 | | | |
| Gln | Arg | Ile | Arg | Arg | Ala | Phe | Val | Ala | Glu | Ala | Gly | Trp | Ala | Leu | Val | | |
| | | 595 | | | | | 600 | | | | | | 605 | | | | |
| Ala | Leu | Asp | Tyr | Ser | Gln | Ile | Glu | Leu | Arg | Val | Leu | Ala | His | Leu | Ser | | |
| | 610 | | | | | 615 | | | | | | 620 | | | | | |
| Gly | Asp | Glu | Asn | Leu | Ile | Arg | Val | Phe | Gln | Glu | Gly | Lys | Asp | Ile | His | | |
| 625 | | | | | 630 | | | | | | 635 | | | | 640 | | |
| Thr | Gln | Thr | Ala | Ser | Trp | Met | Phe | Gly | Val | Pro | Pro | Glu | Ala | Val | Asp | | |
| | | | | 645 | | | | | 650 | | | | | | 655 | | |
| Pro | Leu | Met | Arg | Arg | Ala | Ala | Lys | Thr | Val | Asn | Phe | Gly | Val | Leu | Tyr | | |
| | | | 660 | | | | | 665 | | | | | | 670 | | | |
| Gly | Met | Ser | Ala | His | Arg | Leu | Ser | Gln | Glu | Leu | Ala | Ile | Pro | Tyr | Glu | | |
| | 675 | | | | | | 680 | | | | | 685 | | | | | |
| Glu | Ala | Val | Ala | Phe | Ile | Glu | Arg | Tyr | Phe | Gln | Ser | Phe | Pro | Lys | Val | | |
| | 690 | | | | | 695 | | | | | 700 | | | | | | |
| Arg | Ala | Trp | Ile | Glu | Lys | Thr | Leu | Glu | Glu | Gly | Arg | Lys | Arg | Gly | Tyr | | |
| 705 | | | | | 710 | | | | | 715 | | | | | 720 | | |
| Val | Glu | Thr | Leu | Phe | Gly | Arg | Arg | Arg | Tyr | Val | Pro | Asp | Leu | Asn | Ala | | |
| | | | | 725 | | | | | 730 | | | | | 735 | | | |
| Arg | Val | Lys | Ser | Val | Arg | Glu | Ala | Ala | Glu | Arg | Met | Ala | Phe | Asn | Met | | |
| | | | 740 | | | | | 745 | | | | | | 750 | | | |
| Pro | Val | Gln | Gly | Thr | Ala | Ala | Asp | Leu | Met | Lys | Leu | Ala | Met | Val | Lys | | |
| | | 755 | | | | | 760 | | | | | | 765 | | | | |
| Leu | Phe | Pro | Arg | Leu | Arg | Glu | Met | Gly | Ala | Arg | Met | Leu | Leu | Gln | Val | | |
| | 770 | | | | | | 775 | | | | | 780 | | | | | |
| Ala | Asn | Glu | Leu | Leu | Leu | Glu | Ala | Pro | Gln | Ala | Arg | Ala | Glu | Glu | Val | | |
| 785 | | | | | | 790 | | | | 795 | | | | | 800 | | |

Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Gly His His His His His His
835

<210> 511
<211> 33
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 511
gcctaccgca ccttctttgc cctgaagggc ctc 33

<210> 512
<211> 33
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 512
gaggcccttc agggcaaaga aggtgcggta ggc 33

<210> 513
<211> 2517
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 513
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttocacgcc ctgaagggcc tcaccaccag ccgggggggag 120
ccggtgcagg cggctctacgg cttcgccaag agcctcctca aggccctcag agaggacggg 180
gacgcggtga tcgtggtcct tgacgccgag gccccctcct tccgccacga ggccctacggg 240
gggtacaagg cgggccggggc cccacgcgcg gaggactttc cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggttcacg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480

| | |
|--|------|
| gggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg | 540 |
| gccgactacc gggccctgac cggggacgag tccgacaacc ttcccgggggt caagggcatc | 600 |
| ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag | 660 |
| aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg | 720 |
| aagctctcct gggacctggc caaggtgctc accgacctgc ccctggagggt ggacttcgcc | 780 |
| aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc | 840 |
| agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg | 900 |
| cccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc | 960 |
| gatcttctgg ccctggccgc cgccaggggc ggccgcgtgc accgggcagc agacccttg | 1020 |
| gcggggctaa aggacctcaa ggaggtccgg ggctcctcg ccaaggacct cgccgtcttg | 1080 |
| gcctcgaggg aggggctaga cctcgtgccc ggggacgacc ccatgctcct cgccctacctc | 1140 |
| ctggaccctt cgaacaccac ccccgagggg gtggcgcgcc gctacggggg ggagtggacg | 1200 |
| gaggacgcgc cccaccgggc cctcctctcg gagaggctcc atcggaacct ccttaagcgc | 1260 |
| ctcgaggggg aggagaagct cctttggctc taccacgagg tggaaaagcc cctctcccgg | 1320 |
| gtcctggccc atatggaggc caccggggta cggcgggacg tggcctacct tcaggccctt | 1380 |
| tccctggagc ttgcggagga gatccgcgc ctcgaggagg aggtcttccg cttggcgggc | 1440 |
| cacccttca acctcaactc ccgggaccag ctggaaaggg tgctctttga cgagcttagg | 1500 |
| cttccgcct tgaagaagac gaagaagaca ggcaagcgt ccaccagcgc cgcggtgctg | 1560 |
| gaggccctac gggaggccca ccccatcgtg gagaagatcc tccagcaccg ggagctcacc | 1620 |
| aagctcaaga acacctacgt ggacccccctc ccaagcctcg tccaccgag gacgggccgc | 1680 |
| ctccacaccc gttcaacca gacggccacg gccacgggga ggcttagtag ctccgacccc | 1740 |
| aacctgcaga acatccccgt ccgcaccccc ttggggcaga ggatccgccg ggccttcgtg | 1800 |
| gccgaggcgg gttggcggtt ggtggccctg gactatagcc agatagagct ccgcgtcctc | 1860 |
| gcccacctct ccggggacga aaacctgatc agggctcttc aggaggggaa ggacatccac | 1920 |
| accagaccg caagctggat gttcggcgtc ccccgaggag ccgtggaccc cctgatgcgc | 1980 |
| cggggcgcca agacggtgaa cttcggcgtc ctctacggca tgtccgcca taggctctcc | 2040 |
| caggagcttg ccatccccta cgaggaggcg gtggccttta tagagcgcta cttccaaagc | 2100 |
| ttccccaagg tgcgggcctg gatagaaaag accctggagg aggggaggaa gcggggctac | 2160 |

| | |
|--|------|
| accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag | 480 |
| gggtacctca tcaccccggc ctggctttgg gaaaagtacg gcctgaggcc cgaccagtgg | 540 |
| gccgactacc gggccctgac cggggacgag tccgacaacc ttcccgggggt caagggcatc | 600 |
| ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag | 660 |
| aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggccacat ggacgatctg | 720 |
| aagctctcct gggacctggc caaggtgcg accgacctgc ccctggagggt ggacttcgcc | 780 |
| aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc | 840 |
| agcctcctcc acgagttcgg ctttctggaa agccccaagg ccctggagga gggccctgg | 900 |
| ccccgcggg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc | 960 |
| gatcttctgg ccctggccgc cgccaggggc ggccgcgtgc accgggcagc agaccccttg | 1020 |
| gcggggctaa aggacctcaa ggaggtccgg ggccctcctcg ccaaggacct cgccgtcttg | 1080 |
| gcctcgaggg aggggctaga cctcgtgccc ggggacgacc ccatgctcct cgctacctc | 1140 |
| ctggaccctt cgaacaccac ccccgagggg gtggcgcggc gctacggggg ggagtggacg | 1200 |
| gaggacgcg cccacgggc cctcctctcg gagaggctcc atcggaacct ccttaagcgc | 1260 |
| ctcgaggggg aggagaagct cctttggctc taccacgagg tggaaaagcc cctctcccgg | 1320 |
| gtcctggccc atatggaggc caccggggta cggcgggacg tggcctacct tcaggccctt | 1380 |
| tccctggagc ttgcggagga gatccgcgc ctcgaggagg aggtcttccg cttggcgggc | 1440 |
| caccccttca acctcaactc ccgggaccag ctggaaaggg tgctctttga cgagcttagg | 1500 |
| cttccgcct tgaagaagac gaagaagaca ggcaagcgct ccaccagcgc cgcggtgctg | 1560 |
| gaggccctac gggaggcca ccccatcgtg gagaagatcc tccagcaccg ggagctcacc | 1620 |
| aagctcaaga acacctacgt ggacccctc ccaagcctcg tccacccgag gacgggcccgc | 1680 |
| ctccacacc gttcaacca gacggccacg gccacgggga ggcttagtag ctccgacccc | 1740 |
| aacctgcaga acatccccgt ccgcaccccc ttggggcaga ggatccgcg ggccttcgtg | 1800 |
| gccgaggcgg gttgggcgtt ggtggccctg gactatagcc agatagagct ccgcgtcctc | 1860 |
| gccacctct ccggggacga aaacctgatc agggctcttc aggaggggaa ggacatccac | 1920 |
| accagaccg caagctggat gttcggcgtc ccccgaggc ccgtggaccc cctgatgcgc | 1980 |
| cgggcggcca agacggtgaa cttcggcgtc ctctacggca tgtccgcca taggctctcc | 2040 |
| caggagcttg ccatccccta cgaggaggcg gtggccttta tagagcgcta cttccaaage | 2100 |

```

ttccccaagg tgcgggcctg gatagaaaag accctggagg aggggaggaa gcgggggctac 2160
gtggaaaccc tcttcggaag aaggcgctac gtgcccgacc tcaacgcccg ggtgaagagc 2220
gtcagggagg ccgcgaggcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
ctcatgaagc tcgccatggt gaagctcttc ccccgctcc gggagatggg ggcccgcacg 2340
ctcctccagg tcgccaacga gctcctcctg gaggcccccc aagcgcgggc cgaggagggtg 2400
gcggcttttg ccaaggaggc catggagaag gcctatcccc tcgccgtgcc cctggagggtg 2460
gaggtggggg tgggggagga ctggctttcc gccagggtc accaccacca ccaccac 2517

```

<210> 518
 <211> 839
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 518

```

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1          5          10          15
Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
          20          25          30
Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
          35          40          45
Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
          50          55          60
Val Val Phe Asp Ala Glu Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
65          70          75          80
Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
          85          90          95
Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr Arg Leu
          100          105          110
Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys
          115          120          125
Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
          130          135          140
Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
145          150          155          160
Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
          165          170          175

```

Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
180 185 190

Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
195 200 205

Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
210 215 220

Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
225 230 235 240

Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
245 250 255

Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
260 265 270

Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
275 280 285

Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu
290 295 300

Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala
305 310 315 320

Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala
325 330 335

Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val Arg Gly Leu
340 345 350

Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly Leu Asp Leu
355 360 365

Val Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser
370 375 380

Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr
385 390 395 400

Glu Asp Ala Ala His Arg Ala Leu Leu Ser Glu Arg Leu His Arg Asn
405 410 415

Leu Leu Lys Arg Leu Glu Gly Glu Glu Lys Leu Leu Trp Leu Tyr His
420 425 430

Glu Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met Glu Ala Thr
435 440 445

Gly Val Arg Arg Asp Val Ala Tyr Leu Gln Ala Leu Ser Leu Glu Leu
450 455 460

Ala Glu Glu Ile Arg Arg Leu Glu Glu Glu Val Phe Arg Leu Ala Gly
465 470 475 480

Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala Glu Glu Val
785 790 795 800

Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Gly His His His His His His
835

<210> 519
<211> 33
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 519
gacgacgtcc tggccaccct ggccaagaag gcg 33

<210> 520
<211> 33
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 520
cgctttcttg gccaggggtgg ccaggacgtc gtc 33

<210> 521
<211> 2517
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 521
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
ccggtgcagg cggctctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccgag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggccgggc cccacgccg gaggactttc cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggttcacg cgcctcgagg tcccgggcta cgaggcggac 360

| | |
|---|------|
| gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatcctc | 420 |
| accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag | 480 |
| gggtacctca tcaccccggc ctggccttgg gaaaagtacg gcctgaggcc cgaccagtgg | 540 |
| gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc | 600 |
| ggggagaaga cggcgctcaa gcttctggag gagtggggga gcctggaagc cctcctcaag | 660 |
| aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg | 720 |
| aagctctcct gggacctggc caaggtgcg accgacctgc ccctggagggt ggacttcgcc | 780 |
| aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc | 840 |
| agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg | 900 |
| ccccgcgagg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtggggc | 960 |
| gatcttctgg ccctggccgc cgccaggggc ggccgcgtgc accgggcagc agaccccttg | 1020 |
| gcggggctaa aggacctcaa ggaggtccgg ggctcctcg ccaaggacct cgccgtcttg | 1080 |
| gcctcgaggg aggggctaga cctcgtgccc ggggacgacc ccatgctcct cgcctacctc | 1140 |
| ctggaccctt cgaacaccac ccccgagggg gtggcgcggc gctacggggg ggagtggacg | 1200 |
| gaggacgcgc cccaccgggc cctcctctcg gagaggctcc atcggaacct ccttaagcgc | 1260 |
| ctcgaggggg aggagaagct cctttggctc taccacgagg tggaaaagcc cctctcccgg | 1320 |
| gtcctggccc atatggaggg caccggggta cggcgggacg tggcctacct tcaggccctt | 1380 |
| tccctggagc ttgcggagga gatccgccgc ctcgaggagg aggtcttccg cttggcgggc | 1440 |
| cacccttca acctcaactc ccgggaccag ctggaaaggg tgctctttga cgagcttagg | 1500 |
| cttccgcct tgaagaagac gaagaagaca ggcaagcgct ccaccagcgc cgcggtgctg | 1560 |
| gaggccctac gggaggccca ccccatcggt gagaagatcc tccagcaccg ggagctcacc | 1620 |
| aagctcaaga acacctacgt ggacccccct ccaagcctcg tccacccgag gacgggccgc | 1680 |
| ctccacaccc gcttcaacca gacggccacg gccacgggga ggcttagtag ctccgacccc | 1740 |
| aacctgcaga acatccccgt ccgcaccccc ttggggccaga ggatccgccc ggcttctcgtg | 1800 |
| gccgaggcgg gttgggcgtt ggtggccctg gactatagcc agatagagct ccgcgtcctc | 1860 |
| gccacctct ccggggacga aaacctgatc agggctcttc aggaggggaa ggacatccac | 1920 |
| accagaccg caagctggat gttcggcgtc cccccggagg ccgtggaccc cctgatgcgc | 1980 |
| cggggcgcca agacggtgaa cttcggcgtc ctctacggca tgtccgcca taggctctcc | 2040 |

caggagcttg ccatccccta cgaggaggcg gtggccttta tagagcgcta cttccaaagc 2100
 ttccccaagg tgcgggcctg gatagaaaag accctggagg aggggaggaa gcggggctac 2160
 gtggaaaccc tcttcggaag aaggcgctac gtgcccgacc tcaacgcccg ggtgaagagc 2220
 gtcaggagg cgcgggagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
 ctcatgaagc tcgccatggt gaagctcttc ccccgctcc gggagatggg ggcccgcacg 2340
 ctctccagg tcgccaacga gctctcctg gagggccccc aagcgcgggc cgaggagggtg 2400
 gcggctttgg ccaaggaggc catggagaag gcctatcccc tcgccgtgcc cctggagggtg 2460
 gaggtgggga tgggggagga ctggctttcc gccaaagggtc accaccacca ccaccac 2517

<210> 522
 <211> 839
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 522

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Asn | Ser | Gly | Met | Leu | Pro | Leu | Phe | Glu | Pro | Lys | Gly | Arg | Val | Leu |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Leu | Val | Asp | Gly | His | His | Leu | Ala | Tyr | Arg | Thr | Phe | His | Ala | Leu | Lys |
| | | 20 | | | | | | 25 | | | | | 30 | | |
| Gly | Leu | Thr | Thr | Ser | Arg | Gly | Glu | Pro | Val | Gln | Ala | Val | Tyr | Gly | Phe |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Ala | Lys | Ser | Leu | Leu | Lys | Ala | Leu | Lys | Glu | Asp | Gly | Asp | Ala | Val | Ile |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| Val | Val | Phe | Asp | Ala | Glu | Ala | Pro | Ser | Phe | Arg | His | Glu | Ala | Tyr | Gly |
| 65 | | | | 70 | | | | | 75 | | | | | 80 | |
| Gly | Tyr | Lys | Ala | Gly | Arg | Ala | Pro | Thr | Pro | Glu | Asp | Phe | Pro | Arg | Gln |
| | | | 85 | | | | | 90 | | | | | | 95 | |
| Leu | Ala | Leu | Ile | Lys | Glu | Leu | Val | Asp | Leu | Leu | Gly | Phe | Thr | Arg | Leu |
| | | | 100 | | | | | 105 | | | | | 110 | | |
| Glu | Val | Pro | Gly | Tyr | Glu | Ala | Asp | Asp | Val | Leu | Ala | Ser | Leu | Ala | Lys |
| | | 115 | | | | | 120 | | | | | 125 | | | |
| Lys | Ala | Glu | Lys | Glu | Gly | Tyr | Glu | Val | Arg | Ile | Leu | Thr | Ala | Asp | Lys |
| | 130 | | | | | | 135 | | | | 140 | | | | |
| Asp | Leu | Tyr | Gln | Leu | Leu | Ser | Asp | Arg | Ile | His | Val | Leu | His | Pro | Glu |
| 145 | | | | | 150 | | | | | 155 | | | | | 160 |

Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu Leu Gln Val
770 775 780

Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala Glu Glu Val
785 790 795 800

Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Gly His His His His His His
835

<210> 523
<211> 33
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 523
ggggagaaga cggcgctcaa gcttctggag gag

33

<210> 524
<211> 33
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 524
ctctccaga agcttgagcg ccgtcttctc ccc

33

<210> 525
<211> 2517
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 525
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccgag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggcccgggc cccacgccg gaggactttc cccggcaact cgccctcatc 300

| | | | | | | |
|------------|-------------|-------------|-------------|-------------|-------------|------|
| aaggagctgg | tggacctcct | ggggttcacg | cgcctcgagg | tcccgggcta | cgaggcggac | 360 |
| gacgtcctgg | ccagcctggc | caagaaggcg | gaaaaggagg | gctacgaggt | ccgcacccctc | 420 |
| accgccgaca | aagaccttta | ccagctcctt | tccgaccgca | tccacgtcct | ccacccccgag | 480 |
| gggtacctca | tcacccccgc | ctggcctttg | gaaaagtacg | gcctgaggcc | cgaccagtgg | 540 |
| gccgactacc | gggccctgac | cggggacgag | tccgacaacc | ttcccgggggt | caagggcatc | 600 |
| ggggagaaga | cggcgaggaa | gcttctggag | gagtggggga | gcctggaagc | cctcctcaag | 660 |
| aacctggacc | ggctgaagcc | cgccatccgg | gagaagatcc | tggcccacat | ggacgatctg | 720 |
| aagctctcct | gggacctggc | caaggtgcgc | accgacctgc | ccctggaggt | ggacttcgcc | 780 |
| aaaaggcggg | agccccgaccg | ggaggggctt | aaggcctttc | tggagaggct | tgagtttggc | 840 |
| agcctcctcc | acgagttcgg | ccttctggaa | agccccaagg | ccctggagga | ggccccctgg | 900 |
| ccccgcgagg | aaggggcctt | cgtgggcctt | gtgctttccc | gcaaggagcc | catgtggggc | 960 |
| gatcttctgg | ccctggccgc | cggcaggggc | ggccgcgtgc | accgggcagc | agacccttg | 1020 |
| gcggggctaa | aggacctcaa | ggaggtccgg | ggcctcctcg | ccaaggacct | cgcctcttg | 1080 |
| gcctcgaggg | aggggctaga | cctcgtgccc | ggggacgacc | ccatgctcct | cgcctacctc | 1140 |
| ctggaccctt | cgaacaccac | ccccgagggg | gtggcgcggc | gctacggggg | ggagtggacg | 1200 |
| gaggacgcgc | cccaccgggc | cctcctctcg | gagaggctcc | atcggaacct | ccttaagcgc | 1260 |
| ctcgaggggg | aggagaagct | cctttggctc | taccacgagg | tggaaaagcc | cctctcccg | 1320 |
| gtcctggccc | atatggaggc | caccggggta | cggcgggacg | tggcctacct | tcaggccctt | 1380 |
| tccctggagc | ttgcggagga | gatccgcgcg | ctcgaggagg | aggtcttccg | cttggcgggc | 1440 |
| cacccttca | acctcaactc | ccgggaccag | ctggaaaggg | tgctctttga | cgagcttagg | 1500 |
| cttccgcct | tgaagaagac | gaagaagaca | ggcaagcgct | ccaccagcgc | cgcggtgctg | 1560 |
| gaggccctac | gggaggccca | ccccatcggt | gagaagatcc | tccagcaccg | ggagctcacc | 1620 |
| aagctcaaga | acacctacgt | ggacccccctc | ccaagcctcg | tccaccogag | gacggggcgc | 1680 |
| ctccacaccc | gcttcaacca | gacggccacg | gccacgggga | ggcttagtag | ctccgacccc | 1740 |
| aacctgcaga | acatccccgt | ccgcaccccc | ttggggccaga | ggatccgccg | ggccttcgtg | 1800 |
| gccgaggcgg | gttgggcgtt | gggtggccctg | gactatagcc | agatagagct | ccgcgtcctc | 1860 |
| gccacctct | ccggggacga | aaacctgatc | agggctcttc | aggaggggaa | ggacatccac | 1920 |
| accagaccg | caagctggat | gttcggcgtc | ccccggagg | ccgtggaccc | cctgatgcgc | 1980 |

cgggcggccca agacggtgaa cttcggcgctc ctctacggca tgtccgcccc taggctctcc 2040
caggagcttg ccatcccccta cgaggaggcg gtggccttta tagagcgcta cttccaaagc 2100
ttccccaaagg tgcgggcctg gatagaaaag accctggagg aggggaggaa gcggggctac 2160
gtggaaaccc tcttcggaag aaggcgctac gtgcccgcacc tcaacgcccg ggtgaagagc 2220
gtcaggaggc cgcgggagcg catggccttc aacatgcccg tccagggcac cgccgcccac 2280
ctcatgaagc tcgccatggt gaagctcttc ccccgccctcc gggagatggg ggcccgcacg 2340
ctcctccagg tcgccaacga gctcctcctg gagggccccc aagcgcgggc cgaggagggtg 2400
gcggccttg ccaaggaggc catggagaag gcctatcccc tcgccgtgcc cctggagggtg 2460
gaggtgggga tgggggagga ctggccttcc gccaaagggtc accaccacca ccaccac 2517

<210> 526
<211> 839
<212> PRT
<213> Artificial

<220>
<223> Synthetic

<400> 526

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Asn | Ser | Gly | Met | Leu | Pro | Leu | Phe | Glu | Pro | Lys | Gly | Arg | Val | Leu |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Leu | Val | Asp | Gly | His | His | Leu | Ala | Tyr | Arg | Thr | Phe | His | Ala | Leu | Lys |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Gly | Leu | Thr | Thr | Ser | Arg | Gly | Glu | Pro | Val | Gln | Ala | Val | Tyr | Gly | Phe |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Ala | Lys | Ser | Leu | Leu | Lys | Ala | Leu | Lys | Glu | Asp | Gly | Asp | Ala | Val | Ile |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| Val | Val | Phe | Asp | Ala | Glu | Ala | Pro | Ser | Phe | Arg | His | Glu | Ala | Tyr | Gly |
| 65 | | | | | 70 | | | | | 75 | | | | 80 | |
| Gly | Tyr | Lys | Ala | Gly | Arg | Ala | Pro | Thr | Pro | Glu | Asp | Phe | Pro | Arg | Gln |
| | | | | 85 | | | | | 90 | | | | | 95 | |
| Leu | Ala | Leu | Ile | Lys | Glu | Leu | Val | Asp | Leu | Leu | Gly | Phe | Thr | Arg | Leu |
| | | | 100 | | | | | 105 | | | | | 110 | | |
| Glu | Val | Pro | Gly | Tyr | Glu | Ala | Asp | Asp | Val | Leu | Ala | Ser | Leu | Ala | Lys |
| | | 115 | | | | | 120 | | | | | 125 | | | |
| Lys | Ala | Glu | Lys | Glu | Gly | Tyr | Glu | Val | Arg | Ile | Leu | Thr | Ala | Asp | Lys |
| | 130 | | | | | | 135 | | | | 140 | | | | |
| Asp | Leu | Tyr | Gln | Leu | Leu | Ser | Asp | Arg | Ile | His | Val | Leu | His | Pro | Glu |

| | | | | | | |
|---------------------|---------------------|-----------------|-----------------|-----|--|-----|
| 145 | | 150 | | 155 | | 160 |
| Gly Tyr Leu Ile Thr | Pro Ala Trp Leu Trp | Glu Lys Tyr Gly | Leu Arg | | | |
| | 165 | | 170 | | | 175 |
| Pro Asp Gln Trp | Ala Asp Tyr Arg | Ala Leu Thr Gly | Asp Glu Ser Asp | | | |
| | 180 | 185 | 190 | | | |
| Asn Leu Pro Gly | Val Lys Gly Ile | Gly Glu Lys Thr | Ala Arg Lys Leu | | | |
| | 195 | 200 | 205 | | | |
| Leu Glu Glu Trp | Gly Ser Leu Glu | Ala Leu Leu Lys | Asn Leu Asp Arg | | | |
| | 210 | 215 | 220 | | | |
| Leu Lys Pro Ala | Ile Arg Glu Lys | Ile Leu Ala His | Met Asp Asp Leu | | | |
| | 225 | 230 | 235 | | | 240 |
| Lys Leu Ser Trp | Asp Leu Ala Lys | Val Arg Thr Asp | Leu Pro Leu Glu | | | |
| | 245 | 250 | 255 | | | |
| Val Asp Phe Ala | Lys Arg Arg Glu | Pro Asp Arg Glu | Gly Leu Lys Ala | | | |
| | 260 | 265 | 270 | | | |
| Phe Leu Glu Arg | Leu Glu Phe Gly | Ser Leu Leu His | Glu Phe Gly Leu | | | |
| | 275 | 280 | 285 | | | |
| Leu Glu Ser Pro | Lys Ala Leu Glu | Glu Ala Pro Trp | Pro Pro Pro Glu | | | |
| | 290 | 295 | 300 | | | |
| Gly Ala Phe Val | Gly Phe Val Leu | Ser Arg Lys Glu | Pro Met Trp Ala | | | |
| | 305 | 310 | 315 | | | 320 |
| Asp Leu Leu Ala | Leu Ala Ala Ala | Arg Gly Gly Arg | Val His Arg Ala | | | |
| | 325 | 330 | 335 | | | |
| Ala Asp Pro Leu | Ala Gly Leu Lys | Asp Leu Lys Glu | Val Arg Gly Leu | | | |
| | 340 | 345 | 350 | | | |
| Leu Ala Lys Asp | Leu Ala Val Leu | Ala Ser Arg Glu | Gly Leu Asp Leu | | | |
| | 355 | 360 | 365 | | | |
| Val Pro Gly Asp | Asp Pro Met Leu | Leu Ala Tyr Leu | Leu Asp Pro Ser | | | |
| | 370 | 375 | 380 | | | |
| Asn Thr Thr Pro | Glu Gly Val Ala | Arg Arg Tyr Gly | Gly Glu Trp Thr | | | |
| | 385 | 390 | 395 | | | 400 |
| Glu Asp Ala Ala | His Arg Ala Leu | Leu Ser Glu Arg | Leu His Arg Asn | | | |
| | 405 | 410 | 415 | | | |
| Leu Leu Lys Arg | Leu Glu Gly Glu | Glu Lys Leu Leu | Trp Leu Tyr His | | | |
| | 420 | 425 | 430 | | | |
| Glu Val Glu Lys | Pro Leu Ser Arg | Val Leu Ala His | Met Glu Ala Thr | | | |
| | 435 | 440 | 445 | | | |
| Gly Val Arg Arg | Asp Val Ala Tyr | Leu Gln Ala Leu | Ser Leu Glu Leu | | | |

| | | |
|---|-----|-------------|
| 450 | 455 | 460 |
| Ala Glu Glu Ile Arg Arg Leu Glu Glu Glu Val Phe Arg Leu Ala Gly | | |
| 465 | 470 | 475 480 |
| His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe | | |
| | 485 | 490 495 |
| Asp Glu Leu Arg Leu Pro Ala Leu Lys Lys Thr Lys Lys Thr Gly Lys | | |
| | 500 | 505 510 |
| Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro | | |
| | 515 | 520 525 |
| Ile Val Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys Leu Lys Asn | | |
| | 530 | 535 540 |
| Thr Tyr Val Asp Pro Leu Pro Ser Leu Val His Pro Arg Thr Gly Arg | | |
| | 545 | 550 555 560 |
| Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser | | |
| | 565 | 570 575 |
| Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly | | |
| | 580 | 585 590 |
| Gln Arg Ile Arg Arg Ala Phe Val Ala Glu Ala Gly Trp Ala Leu Val | | |
| | 595 | 600 605 |
| Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser | | |
| | 610 | 615 620 |
| Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys Asp Ile His | | |
| | 625 | 630 635 640 |
| Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu Ala Val Asp | | |
| | 645 | 650 655 |
| Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly Val Leu Tyr | | |
| | 660 | 665 670 |
| Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu | | |
| | 675 | 680 685 |
| Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val | | |
| | 690 | 695 700 |
| Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys Arg Gly Tyr | | |
| | 705 | 710 715 720 |
| Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn Ala | | |
| | 725 | 730 735 |
| Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met | | |
| | 740 | 745 750 |
| Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys | | |

Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
 145 150 155 160
 Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
 165 170 175
 Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
 180 185 190
 Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
 195 200 205
 Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
 210 215 220
 Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
 225 230 235 240
 Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
 245 250 255
 Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
 260 265 270
 Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
 275 280 285
 Leu Gly Gly Glu Lys Pro Arg Glu Glu Ala Pro Trp Pro Pro Pro Glu
 290 295 300
 Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala
 305 310 315 320
 Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala
 325 330 335
 Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val Arg Gly Leu
 340 345 350
 Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly Leu Asp Leu
 355 360 365
 Val Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser
 370 375 380
 Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr
 385 390 395 400
 Glu Asp Ala Ala His Arg Ala Leu Leu Ser Glu Arg Leu His Arg Asn
 405 410 415
 Leu Leu Lys Arg Leu Glu Gly Glu Glu Lys Leu Leu Trp Leu Tyr His
 420 425 430
 Glu Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met Glu Ala Thr
 435 440 445

| | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Gly | Val | Arg | Arg | Asp | Val | Ala | Tyr | Leu | Gln | Ala | Leu | Ser | Leu | Glu | Leu | 450 | 455 | 460 |
| Ala | Glu | Glu | Ile | Arg | Arg | Leu | Glu | Glu | Glu | Val | Phe | Arg | Leu | Ala | Gly | 465 | 470 | 475 |
| His | Pro | Phe | Asn | Leu | Asn | Ser | Arg | Asp | Gln | Leu | Glu | Arg | Val | Leu | Phe | 485 | 490 | 495 |
| Asp | Glu | Leu | Arg | Leu | Pro | Ala | Leu | Lys | Lys | Thr | Lys | Lys | Thr | Gly | Lys | 500 | 505 | 510 |
| Arg | Ser | Thr | Ser | Ala | Ala | Val | Leu | Glu | Ala | Leu | Arg | Glu | Ala | His | Pro | 515 | 520 | 525 |
| Ile | Val | Glu | Lys | Ile | Leu | Gln | His | Arg | Glu | Leu | Thr | Lys | Leu | Lys | Asn | 530 | 535 | 540 |
| Thr | Tyr | Val | Asp | Pro | Leu | Pro | Ser | Leu | Val | His | Pro | Arg | Thr | Gly | Arg | 545 | 550 | 555 |
| Leu | His | Thr | Arg | Phe | Asn | Gln | Thr | Ala | Thr | Ala | Thr | Gly | Arg | Leu | Ser | 565 | 570 | 575 |
| Ser | Ser | Asp | Pro | Asn | Leu | Gln | Asn | Ile | Pro | Val | Arg | Thr | Pro | Leu | Gly | 580 | 585 | 590 |
| Gln | Arg | Ile | Arg | Arg | Ala | Phe | Val | Ala | Glu | Ala | Gly | Trp | Ala | Leu | Val | 595 | 600 | 605 |
| Ala | Leu | Asp | Tyr | Ser | Gln | Ile | Glu | Leu | Arg | Val | Leu | Ala | His | Leu | Ser | 610 | 615 | 620 |
| Gly | Asp | Glu | Asn | Leu | Ile | Arg | Val | Phe | Gln | Glu | Gly | Lys | Asp | Ile | His | 625 | 630 | 635 |
| Thr | Gln | Thr | Ala | Ser | Trp | Met | Phe | Gly | Val | Pro | Pro | Glu | Ala | Val | Asp | 645 | 650 | 655 |
| Pro | Leu | Met | Arg | Arg | Ala | Ala | Lys | Thr | Val | Asn | Phe | Gly | Val | Leu | Tyr | 660 | 665 | 670 |
| Gly | Met | Ser | Ala | His | Arg | Leu | Ser | Gln | Glu | Leu | Ala | Ile | Pro | Tyr | Glu | 675 | 680 | 685 |
| Glu | Ala | Val | Ala | Phe | Ile | Glu | Arg | Tyr | Phe | Gln | Ser | Phe | Pro | Lys | Val | 690 | 695 | 700 |
| Arg | Ala | Trp | Ile | Glu | Lys | Thr | Leu | Glu | Glu | Gly | Arg | Lys | Arg | Gly | Tyr | 705 | 710 | 715 |
| Val | Glu | Thr | Leu | Phe | Gly | Arg | Arg | Arg | Tyr | Val | Pro | Asp | Leu | Asn | Ala | 725 | 730 | 735 |
| Arg | Val | Lys | Ser | Val | Arg | Glu | Ala | Ala | Glu | Arg | Met | Ala | Phe | Asn | Met | 740 | 745 | 750 |

Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
755 760 765

Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu Leu Gln Val
770 775 780

Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala Glu Glu Val
785 790 795 800

Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Gly His His His His His His
835

<210> 531
<211> 54
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 531
cacgagttcg gccttctggg aggggagaag ccccgaggagg agggccccctg gccc 54

<210> 532
<211> 54
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 532
gggccagggg gcctcctccc ggggcttctc cctcccaga aggccgaact cgtg 54

<210> 533
<211> 2517
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 533
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60

caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120

ccggtgcagg cggctctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180

| | |
|--|------|
| gacgcggtga tcgtggtctt tgacgccgag gccccctcct tccgccacga ggcctacggg | 240 |
| gggtacaagg cgggccgggc cccacgccg gaggactttc cccggcaact cgcctcatc | 300 |
| aaggagctgg tggacctcct ggggttcacg cgcctcgagg tcccgggcta cgaggcggac | 360 |
| gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatactc | 420 |
| accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag | 480 |
| gggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg | 540 |
| gccgactacc gggccctgac cggggacgag tccgacaacc tcccgggggt caagggcatc | 600 |
| ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag | 660 |
| aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccatat ggacgatctg | 720 |
| aagctctcct gggacctggc caagggtgcg accgacctgc ccctggagggt ggacttcgcc | 780 |
| aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc | 840 |
| agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg | 900 |
| ccccgcggg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc | 960 |
| gatcttctgg ccctggccgc ctgcaggggc ggccgcgtgc accgggcagc agacccttg | 1020 |
| gcggggctaa aggacctcaa ggaggtccgg ggctcctcg ccaaggacct cgcgtcttg | 1080 |
| gcctcgaggg aggggctaga cctcgtgccc ggggacgacc ccatgctcct cgcctacctc | 1140 |
| ctggaccctt cgaacaccac ccccgagggg gtggcgcggc gctacggggg ggagtggacg | 1200 |
| gaggacgccg cccaccgggc cctcctctcg gagaggctcc atcggaacct ccttaagcgc | 1260 |
| ctcgaggggg aggagaagct cctttggctc taccacgagg tggaaaagcc cctctcccgg | 1320 |
| gtcctggccc atatggaggc caccggggta cggcgggacg tggcctacct tcaggccctt | 1380 |
| tccctggagc ttgcggagga gatccgccgc ctcgaggagg aggtcttccg cttggcgggc | 1440 |
| cacccttca acctcaactc ccgggaccag ctggaaaggg tgctctttga cgagcttagg | 1500 |
| cttccgcct tgaagaagac gaagaagaca ggcaagcgt ccaccagcgc cgcggtgctg | 1560 |
| gaggccctac gggaggccca ccccatcgtg gagaagatcc tccagcaccg ggagctcacc | 1620 |
| aagctcaaga acacctacgt ggacccccct ccaagcctcg tccaccgag gacgggcgc | 1680 |
| ctccacaccc gcttcaacca gacggccacg gccacgggga ggcttagtag ctccgacccc | 1740 |
| aacctgcaga acatccccgt ccgcaccccc ttgggcccaga ggatccgccg ggcttctgtg | 1800 |
| gccgaggcgg gttgggcgtt ggtggccctg gactatagcc agatagagct ccgcgtcctc | 1860 |

gcccacctct ccggggacga aaacctgata agggctcttc aggaggggaa ggacatccac 1920
 acccagaccg caagctggat gttcggcgtc ccccgagg cctgggaccc cctgatgcgc 1980
 cgggaggcca agacggtgaa cttcggcgtc ctctacggca tgtccgcca taggctctcc 2040
 caggagcttg ccaccccta cgaggaggcg gtggccttta tagagcgcta cttccaaagc 2100
 ttccccaagg tgcgggcctg gatagaaaag accctggagg aggggaggaa gcggggctac 2160
 gtggaaaccc tcttcggaag aaggcgctac gtgcccgcacc tcaacgcccg ggtgaagagc 2220
 gtcaggagg cgcggagcg catggccttc aacatgcccg tccagggcac cgcgcgcgac 2280
 ctcatgaagc tcgcatggt gaagctcttc cccgcctcc gggagatggg ggcccgcgac 2340
 ctctccagg tcgccaacga gctcctctg gaggccccc aagcgcgggc cgaggagggtg 2400
 gcggctttgg ccaaggaggc catggagaag gcctatcccc tcgccgtgcc cctggagggtg 2460
 gaggtgggga tgggggagga ctggctttcc gccagggtc accaccacca ccaccac 2517

<210> 534
 <211> 839
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 534

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Asn | Ser | Gly | Met | Leu | Pro | Leu | Phe | Glu | Pro | Lys | Gly | Arg | Val | Leu |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Leu | Val | Asp | Gly | His | His | Leu | Ala | Tyr | Arg | Thr | Phe | His | Ala | Leu | Lys |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Gly | Leu | Thr | Thr | Ser | Arg | Gly | Glu | Pro | Val | Gln | Ala | Val | Tyr | Gly | Phe |
| | | | 35 | | | | 40 | | | | | 45 | | | |
| Ala | Lys | Ser | Leu | Leu | Lys | Ala | Leu | Lys | Glu | Asp | Gly | Asp | Ala | Val | Ile |
| | | | 50 | | | 55 | | | | | 60 | | | | |
| Val | Val | Phe | Asp | Ala | Glu | Ala | Pro | Ser | Phe | Arg | His | Glu | Ala | Tyr | Gly |
| 65 | | | | | 70 | | | | | 75 | | | | 80 | |
| Gly | Tyr | Lys | Ala | Gly | Arg | Ala | Pro | Thr | Pro | Glu | Asp | Phe | Pro | Arg | Gln |
| | | | | 85 | | | | | 90 | | | | | 95 | |
| Leu | Ala | Leu | Ile | Lys | Glu | Leu | Val | Asp | Leu | Leu | Gly | Phe | Thr | Arg | Leu |
| | | | 100 | | | | | 105 | | | | | 110 | | |
| Glu | Val | Pro | Gly | Tyr | Glu | Ala | Asp | Asp | Val | Leu | Ala | Ser | Leu | Ala | Lys |
| | | | 115 | | | | 120 | | | | | 125 | | | |

| | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|--|
| Lys | Ala | Glu | Lys | Glu | Gly | Tyr | Glu | Val | Arg | Ile | Leu | Thr | Ala | Asp | Lys | | |
| 130 | | | | | | 135 | | | | | 140 | | | | | | |
| Asp | Leu | Tyr | Gln | Leu | Leu | Ser | Asp | Arg | Ile | His | Val | Leu | His | Pro | Glu | | |
| 145 | | | | 150 | | | | | | 155 | | | | | 160 | | |
| Gly | Tyr | Leu | Ile | Thr | Pro | Ala | Trp | Leu | Trp | Glu | Lys | Tyr | Gly | Leu | Arg | | |
| | | | | 165 | | | | | 170 | | | | | 175 | | | |
| Pro | Asp | Gln | Trp | Ala | Asp | Tyr | Arg | Ala | Leu | Thr | Gly | Asp | Glu | Ser | Asp | | |
| | | | 180 | | | | | 185 | | | | | 190 | | | | |
| Asn | Leu | Pro | Gly | Val | Lys | Gly | Ile | Gly | Glu | Lys | Thr | Ala | Arg | Lys | Leu | | |
| | | 195 | | | | | 200 | | | | | 205 | | | | | |
| Leu | Glu | Glu | Trp | Gly | Ser | Leu | Glu | Ala | Leu | Leu | Lys | Asn | Leu | Asp | Arg | | |
| | 210 | | | | | 215 | | | | | 220 | | | | | | |
| Leu | Lys | Pro | Ala | Ile | Arg | Glu | Lys | Ile | Leu | Ala | His | Met | Asp | Asp | Leu | | |
| 225 | | | | | 230 | | | | | 235 | | | | | 240 | | |
| Lys | Leu | Ser | Trp | Asp | Leu | Ala | Lys | Val | Arg | Thr | Asp | Leu | Pro | Leu | Glu | | |
| | | | | 245 | | | | | 250 | | | | | 255 | | | |
| Val | Asp | Phe | Ala | Lys | Arg | Arg | Glu | Pro | Asp | Arg | Glu | Arg | Leu | Arg | Ala | | |
| | | | 260 | | | | | 265 | | | | | 270 | | | | |
| Phe | Leu | Glu | Arg | Leu | Glu | Phe | Gly | Ser | Leu | Leu | His | Glu | Phe | Gly | Leu | | |
| | | 275 | | | | | 280 | | | | | 285 | | | | | |
| Leu | Glu | Ser | Pro | Lys | Ala | Leu | Glu | Glu | Ala | Pro | Trp | Pro | Pro | Pro | Glu | | |
| | 290 | | | | | 295 | | | | | 300 | | | | | | |
| Gly | Ala | Phe | Val | Gly | Phe | Val | Leu | Ser | Arg | Lys | Glu | Pro | Met | Trp | Ala | | |
| 305 | | | | | 310 | | | | | 315 | | | | | 320 | | |
| Asp | Leu | Leu | Ala | Leu | Ala | Ala | Cys | Arg | Gly | Gly | Arg | Val | His | Arg | Ala | | |
| | | | | 325 | | | | | 330 | | | | | 335 | | | |
| Ala | Asp | Pro | Leu | Ala | Gly | Leu | Lys | Asp | Leu | Lys | Glu | Val | Arg | Gly | Leu | | |
| | | | 340 | | | | | 345 | | | | | 350 | | | | |
| Leu | Ala | Lys | Asp | Leu | Ala | Val | Leu | Ala | Ser | Arg | Glu | Gly | Leu | Asp | Leu | | |
| | | 355 | | | | | 360 | | | | | 365 | | | | | |
| Val | Pro | Gly | Asp | Asp | Pro | Met | Leu | Leu | Ala | Tyr | Leu | Leu | Asp | Pro | Ser | | |
| | 370 | | | | | 375 | | | | | 380 | | | | | | |
| Asn | Thr | Thr | Pro | Glu | Gly | Val | Ala | Arg | Arg | Tyr | Gly | Gly | Glu | Trp | Thr | | |
| 385 | | | | | 390 | | | | | 395 | | | | | 400 | | |
| Glu | Asp | Ala | Ala | His | Arg | Ala | Leu | Leu | Ser | Glu | Arg | Leu | His | Arg | Asn | | |
| | | | | 405 | | | | | 410 | | | | | 415 | | | |
| Leu | Leu | Lys | Arg | Leu | Glu | Gly | Glu | Glu | Lys | Leu | Leu | Trp | Leu | Tyr | His | | |
| | | | 420 | | | | | 425 | | | | | 430 | | | | |

Glu Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met Glu Ala Thr
 435 440 445
 Gly Val Arg Arg Asp Val Ala Tyr Leu Gln Ala Leu Ser Leu Glu Leu
 450 455 460
 Ala Glu Glu Ile Arg Arg Leu Glu Glu Glu Val Phe Arg Leu Ala Gly
 465 470 475 480
 His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe
 485 490 495
 Asp Glu Leu Arg Leu Pro Ala Leu Lys Lys Thr Lys Lys Thr Gly Lys
 500 505 510
 Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro
 515 520 525
 Ile Val Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys Leu Lys Asn
 530 535 540
 Thr Tyr Val Asp Pro Leu Pro Ser Leu Val His Pro Arg Thr Gly Arg
 545 550 555 560
 Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser
 565 570 575
 Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly
 580 585 590
 Gln Arg Ile Arg Arg Ala Phe Val Ala Glu Ala Gly Trp Ala Leu Val
 595 600 605
 Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser
 610 615 620
 Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys Asp Ile His
 625 630 635 640
 Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu Ala Val Asp
 645 650 655
 Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly Val Leu Tyr
 660 665 670
 Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
 675 680 685
 Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
 690 695 700
 Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys Arg Gly Tyr
 705 710 715 720
 Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn Ala
 725 730 735

| | |
|--|------|
| ccggtgcagg cggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg | 180 |
| gacgcggtga tcgtggtctt tgacgccgag gccccctcct tccgccacga ggcctacggg | 240 |
| gggtacaagg cgggccgggc cccacgcgcg gaggactttc cccggcaact cgcctcatc | 300 |
| aaggagctgg tggacctcct ggggttcacg cgcctcgagg tcccgggcta cgaggcggac | 360 |
| gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatactc | 420 |
| accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag | 480 |
| gggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg | 540 |
| gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc | 600 |
| ggggagaaga cggcgaggaa gcttctgaag gagtggggga gcctggaagc cctcctcaag | 660 |
| aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccatat ggacgatctg | 720 |
| aagctctcct gggacctggc caaggtgcmc accgacctgc ccctggagggt ggacttcgcc | 780 |
| aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc | 840 |
| agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg | 900 |
| ccccgcggg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtggggc | 960 |
| gatcttctgg ccctggccgc cgccaggggc ggccgcgtgc accgggcagc agaccccttg | 1020 |
| gcggggctaa aggacctcaa ggagggtccg ggccctctcg ccaaggacct cgccgtcttg | 1080 |
| gcctcgaggg aggggctaga cctcgtgccc ggggacgacc ccatgctcct cgctacctc | 1140 |
| ctggaccctt cgaacaccac ccccgagggg gtggcgcggc gctacggggg ggagtggacg | 1200 |
| gaggacgccg cccaccgggc cctcctctcg gagaggctcc atcggaacct ccttaagcmc | 1260 |
| ctcgaggggg aggagaagct cctttggctc taccacgagg tggaaaagcc cctctcccgg | 1320 |
| gtcctggccc atatggaggc caccggggta cggcgggacg tggcctacct tcaggccctt | 1380 |
| tccttgagc ttgcggagga gatccgccgc ctcgaggagg aggtcttccg cttggcgggc | 1440 |
| cacccttca acctcaactc ccgggaccag ctggaaaggg tgctctttga cgagcttagg | 1500 |
| cttccgcct tgaagaagac gaagaagaca ggcaagcgt ccaccagcmc cgcggtgctg | 1560 |
| gaggccctac gggaggccca ccccatcgtg gagaagatcc tccagcaccg ggagctcacc | 1620 |
| aagctcaaga acacctacgt ggacccccct ccaagcctcg tccacccgag gacgggccgc | 1680 |
| ctccacaccc gcttcaacca gacggccacg gccacgggga ggcttagtag ctccgacccc | 1740 |
| aacctgcaga acatccccgt ccgcaccccc ttgggccaga ggatccgccg ggccttcgtg | 1800 |

| | | | | | | |
|-------------|-------------|-------------|------------|------------|-------------|------|
| ggccaccacc | tggcctaccg | caccttcttt | gccctgaagg | gcctcaccac | cagccggggg | 120 |
| gagccggtgc | aggcggtcta | cggcttcgcc | aagagcctcc | tcaaggccct | cagagaggac | 180 |
| ggggacgcgg | tgatcgtggt | ctttgacgcc | gaggccccct | ccttccgcca | cgaggcctac | 240 |
| ggggggtaca | aggcgggccg | ggccccacg | ccggaggact | ttccccggca | actcgccctc | 300 |
| atcaaggagc | tgggtggacct | cctgggggttc | acgcgcctcg | aggtcccggg | ctacgaggcg | 360 |
| gacgacgtcc | tggccaccct | ggccaagaag | gcggaagagg | agggctacga | ggtccgcata | 420 |
| ctcaccgccc | acaaagacct | ttaccagctc | ctttccgacc | gcatccacgt | cctccacccc | 480 |
| gaggggtacc | tcatcacccc | ggcctggctt | tgggaaaagt | acggcctgag | gcccgaccag | 540 |
| tgggcccact | accggggcct | gaccggggac | gagtccgaca | accttcccgg | ggtcaagggc | 600 |
| atcgggggaga | agacggcgct | caagcttctg | gaggagtggg | ggagcctgga | agccctcctc | 660 |
| aagaacctgg | accggctgaa | gcccgccatc | cgggagaaga | tcctggccca | catggacgat | 720 |
| ctgaagctct | cctgggacct | ggccaaggtg | cgcaccgacc | tgccctgga | ggtggacttc | 780 |
| gccaaaaggc | gggagcccga | ccgggagggg | cttaaggcct | ttctggagag | gcttgagttt | 840 |
| ggcagcctcc | tccacgagtt | cggccttctg | ggaggggaga | agccccggga | ggaggccccc | 900 |
| tggcccccg | cgaaggggc | cttcgtgggc | tttgtgtttt | cccgcaagga | gcccattgtg | 960 |
| gcccattctc | tggccctggc | cgcctgcagg | ggcgcccgcg | tgcaccgggc | agcagacccc | 1020 |
| ttggcggggc | taaaggacct | caaggaggtc | cggggcctcc | tcgccaagga | cctcgccgtc | 1080 |
| ttggcctcga | gggaggggct | agacctcgtg | cccggggacg | accccatgct | cctcgccctac | 1140 |
| ctcctggacc | cttcgaacac | cacccccgag | ggggtggcgc | ggcgctacgg | gggggagtgg | 1200 |
| acggaggacg | ccgcccaccg | ggccctcctc | tcggagaggc | tccatcgga | cctccttaag | 1260 |
| cgcctcgagg | gggaggagaa | gctccttttg | ctctaccacg | aggtggaaaa | gcccctctcc | 1320 |
| cgggtcctgg | cccatatgga | ggccaccggg | gtacggcggg | acgtggccta | ccttcaggcc | 1380 |
| ctttccctgg | agcttgcgga | ggagatccgc | cgcctcgagg | aggaggtctt | ccgcttgcg | 1440 |
| ggccaccctt | tcaacctcaa | ctcccgggac | cagctggaaa | gggtgctctt | tgacgagctt | 1500 |
| aggcttcccc | ccttgaagaa | gacgaagaag | acaggcaagc | gctccaccag | cgccgcggtg | 1560 |
| ctggaggccc | tacgggaggc | ccaccccatc | gtggagaaga | tcctccagca | ccgggagctc | 1620 |
| accaagctca | agaacaccta | cgtggacccc | ctcccaagcc | tcgtccaccc | gaggacgggc | 1680 |
| cgcctccaca | cccgttcaa | ccagacggcc | acggccacgg | ggaggcttag | tagctccgac | 1740 |

```

cccaacctgc agaacatccc cgtccgcacc cccttggggc agaggatccg ccgggccttc 1800
gtggccgagg cgggttgggc gttggtggcc ctggactata gccagataga gctccgcgtc 1860
ctcgcaccacc tctccgggga cgaaaacctg atcaggggtct tocaggaggg gaaggacatc 1920
cacaccacaga ccgcaagctg gatgttcggc gtcccccccg aggccgtgga cccctgatg 1980
cgccggggcg ccaagacggt gaacttcggc gtcctctacg gcatgtccgc ccataggctc 2040
tcccaggagc ttgccatccc ctacgaggag gcggtggcct ttatagagcg ctacttccaa 2100
agcttcccca aggtgcgggc ctggatagaa aagaccctgg aggaggggag gaagcggggc 2160
tacgtggaaa ccctcttcgg aagaaggcgc tacgtgcccg acctcaacgc ccgggtgaag 2220
agcgtcaggg aggccgcgga gcgcatggcc ttcaacatgc ccgtccaggg caccgccgcc 2280
gacctcatga agctcgccat ggtgaagctc ttcccccgcc tccgggagat gggggcccgc 2340
atgctcctcc aggtcgccaa cgagctctc ctggaggccc cccaagcgcg ggccgaggag 2400
gtggcggtt tggccaagga ggccatggag aaggcctatc ccctcgccgt gcccctggag 2460
gtggaggtgg ggatggggga ggactggctt tccgccaagg gtcaccacca ccaccaccac 2520

```

```

<210> 542
<211> 840
<212> PRT
<213> Artificial

```

```

<220>
<223> Synthetic

```

```

<400> 542

```

```

Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
1           5           10           15

Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu
          20           25           30

Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
          35           40           45

Phe Ala Lys Ser Leu Leu Lys Ala Leu Arg Glu Asp Gly Asp Ala Val
          50           55           60

Ile Val Val Phe Asp Ala Glu Ala Pro Ser Phe Arg His Glu Ala Tyr
65           70           75           80

Gly Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg
          85           90           95

Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr Arg
          100          105          110

```


Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn
725 730 735

Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn
740 745 750

Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val
755 760 765

Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu Leu Gln
770 775 780

Val Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala Glu Glu
785 790 795 800

Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro Leu Ala
805 810 815

Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala
820 825 830

Lys Gly His His His His His His
835 840

<210> 543
<211> 2520
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 543
atgaattcgg aggcgatgct gcccctcttt gagcccaagg gccgggtcct cctggtggac 60
ggccaccacc tggcctaccg caccttcttt gccctgaagg gcctcaccac cagccggggg 120
gagccggtgc aggcggtcta cggttcgcc aagagcctcc tcaaggccct cagagaggac 180
ggggacgcgg tgatcgtggt ctttgacgcc gaggccccct ccttccgcca cgaggcctac 240
gggggggtaca aggcgggccc ggccccacg ccggaggact ttccccggca actcgcctc 300
atcaaggagc tgggtggacct cctggggttc acgcgcctcg aggtcccggg ctacgaggcg 360
gacgacgtcc tggccaccct ggccaagaag gcggaaaagg agggctacga ggtccgcac 420
ctcaccgccg acaaagacct ttaccagctc ctttccgacc gcatccacgt cctccacccc 480
gaggggtacc tcatcaccoc ggccctggctt tgggaaaagt acggcctgag gcccgaccag 540
tgggccgact accgggccct gaccggggac gagtccgaca accttcccgg ggtcaagggc 600
atcggggaga agacggcgct caagcttctg aaggagtggg ggagcctgga agccctcctc 660
aagaacctgg accggctgaa gcccgccatc cgggagaaga tcctggccca catggacgat 720

| | |
|---|------|
| ctgaagctct cctgggacct ggccaagggtg cgcaccgacc tgccccctgga ggtggacttc | 780 |
| gccaaaaggc gggagcccgga ccgggagggg cttaaggcct ttctggagag gcttgagttt | 840 |
| ggcagcctcc tccacgagtt cggccttctg ggaggggaga agccccggga ggaggccccc | 900 |
| tggcccccg cgggaaggggc cttcgtgggc tttgtgcttt cccgcaagga gcccatgtgg | 960 |
| gccgatcttc tggccctggc cgctgcagg ggcggccgcg tgcaccgggc agcagacccc | 1020 |
| ttggcggggc taaaggacct caaggaggtc cggggcctcc tcgccaagga cctcgccgtc | 1080 |
| ttggcctcga gggaggggct agacctcgtg cccggggacg accccatgct cctcgccctac | 1140 |
| ctcctggacc cttcgaacac ccccccgag ggggtggcgc ggcgctacgg gggggagtgg | 1200 |
| acggaggacg ccgcccaccg ggccctctc tcggagaggc tccatcgga cctccttaag | 1260 |
| cgctcgagg gggaggagaa gctcctttgg ctctaccacg aggtggaaaa gccctctcc | 1320 |
| cgggtcctgg cccatatgga ggccaccggg gtacggcggg acgtggccta ccttcaggcc | 1380 |
| ctttccctgg agcttgcgga ggagatccgc cgctcgagg aggaggtctt ccgcttgcg | 1440 |
| ggccaccct tcaacctcaa ctcccgggac cagctggaaa gggtgctctt tgacgagctt | 1500 |
| aggcttcccg ccttgaagaa gacgaagaag acaggcaagc gctccaccag cgccgcggtg | 1560 |
| ctggaggccc tacgggaggc ccaccccatc gtggagaaga tcctccagca ccgggagctc | 1620 |
| accaagctca agaacaccta cgtggacccc ctcccaagcc tcgtccacc gaggacgggc | 1680 |
| cgctccaca cccgcttcaa ccagacggcc acggccacgg ggaggcttag tagctccgac | 1740 |
| cccaacctgc agaacatccc cgtccgcacc cccttgggccc agaggatccg ccgggccttc | 1800 |
| gtggccgagg cgggttgggc gttggtggcc ctggactata gccagataga gctccgcgtc | 1860 |
| ctcgcccacc tctccgggga cgaaaacctg atcagggtct tccaggagg gaaggacatc | 1920 |
| cacaccaga ccgcaagctg gatgttcggc gtccccccgg aggccgtgga cccctgatg | 1980 |
| cgccgggcgg ccaagacggt gaacttcggc gtctctacg gcatgtccgc ccataggctc | 2040 |
| tcccaggagc ttgccatccc ctacgaggag gcggtggcct ttatagagcg ctacttcaa | 2100 |
| agcttcccca aggtgcgggc ctggatagaa aagaccctgg aggaggggag gaagcggggc | 2160 |
| tacgtgaaa cctcttcgg aagaaggcgc tacgtgcccg acctcaacgc ccgggtgaag | 2220 |
| aggccgcgga gcgcatggcc ttcaacatgc ccgtccaggg caccgccgcc | 2280 |
| gacctcatga agctcgccat ggtgaagctc tcccccgcc tccgggagat gggggcccg | 2340 |
| atgctcctcc aggtcgccaa cgagctctc ctggaggccc cccaagcgc ggccgaggag | 2400 |
| gtggcggtt tggccaagga ggccatggag aaggcctatc ccctcgccgt gccctggag | 2460 |

gtggaggtgg ggatggggga ggactggctt tccgccaagg gtcaccacca ccaccaccac 2520

<210> 544
 <211> 840
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 544

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Asn | Ser | Glu | Ala | Met | Leu | Pro | Leu | Phe | Glu | Pro | Lys | Gly | Arg | Val | 1 | 5 | 10 | 15 |
| Leu | Leu | Val | Asp | Gly | His | His | Leu | Ala | Tyr | Arg | Thr | Phe | Phe | Ala | Leu | 20 | 25 | 30 | |
| Lys | Gly | Leu | Thr | Thr | Ser | Arg | Gly | Glu | Pro | Val | Gln | Ala | Val | Tyr | Gly | 35 | 40 | 45 | |
| Phe | Ala | Lys | Ser | Leu | Leu | Lys | Ala | Leu | Arg | Glu | Asp | Gly | Asp | Ala | Val | 50 | 55 | 60 | |
| Ile | Val | Val | Phe | Asp | Ala | Glu | Ala | Pro | Ser | Phe | Arg | His | Glu | Ala | Tyr | 65 | 70 | 75 | 80 |
| Gly | Gly | Tyr | Lys | Ala | Gly | Arg | Ala | Pro | Thr | Pro | Glu | Asp | Phe | Pro | Arg | 85 | 90 | 95 | |
| Gln | Leu | Ala | Leu | Ile | Lys | Glu | Leu | Val | Asp | Leu | Leu | Gly | Phe | Thr | Arg | 100 | 105 | 110 | |
| Leu | Glu | Val | Pro | Gly | Tyr | Glu | Ala | Asp | Asp | Val | Leu | Ala | Thr | Leu | Ala | 115 | 120 | 125 | |
| Lys | Lys | Ala | Glu | Lys | Glu | Gly | Tyr | Glu | Val | Arg | Ile | Leu | Thr | Ala | Asp | 130 | 135 | 140 | |
| Lys | Asp | Leu | Tyr | Gln | Leu | Leu | Ser | Asp | Arg | Ile | His | Val | Leu | His | Pro | 145 | 150 | 155 | 160 |
| Glu | Gly | Tyr | Leu | Ile | Thr | Pro | Ala | Trp | Leu | Trp | Glu | Lys | Tyr | Gly | Leu | 165 | 170 | 175 | |
| Arg | Pro | Asp | Gln | Trp | Ala | Asp | Tyr | Arg | Ala | Leu | Thr | Gly | Asp | Glu | Ser | 180 | 185 | 190 | |
| Asp | Asn | Leu | Pro | Gly | Val | Lys | Gly | Ile | Gly | Glu | Lys | Thr | Ala | Leu | Lys | 195 | 200 | 205 | |
| Leu | Leu | Lys | Glu | Trp | Gly | Ser | Leu | Glu | Ala | Leu | Leu | Lys | Asn | Leu | Asp | 210 | 215 | 220 | |
| Arg | Leu | Lys | Pro | Ala | Ile | Arg | Glu | Lys | Ile | Leu | Ala | His | Met | Asp | Asp | | | | |

| | | | | | | |
|---|--|-----|--|-----|--|-----|
| 225 | | 230 | | 235 | | 240 |
| Leu Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu | | | | | | |
| | | 245 | | 250 | | 255 |
| Glu Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Gly Leu Lys | | | | | | |
| | | 260 | | 265 | | 270 |
| Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly | | | | | | |
| | | 275 | | 280 | | 285 |
| Leu Leu Gly Gly Glu Lys Pro Arg Glu Glu Ala Pro Trp Pro Pro Pro | | | | | | |
| | | 290 | | 295 | | 300 |
| Glu Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp | | | | | | |
| | | 305 | | 310 | | 315 |
| Ala Asp Leu Leu Ala Leu Ala Ala Cys Arg Gly Gly Arg Val His Arg | | | | | | |
| | | 325 | | 330 | | 335 |
| Ala Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val Arg Gly | | | | | | |
| | | 340 | | 345 | | 350 |
| Leu Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly Leu Asp | | | | | | |
| | | 355 | | 360 | | 365 |
| Leu Val Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro | | | | | | |
| | | 370 | | 375 | | 380 |
| Ser Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp | | | | | | |
| | | 385 | | 390 | | 395 |
| Thr Glu Asp Ala Ala His Arg Ala Leu Leu Ser Glu Arg Leu His Arg | | | | | | |
| | | 405 | | 410 | | 415 |
| Asn Leu Leu Lys Arg Leu Glu Gly Glu Glu Lys Leu Leu Trp Leu Tyr | | | | | | |
| | | 420 | | 425 | | 430 |
| His Glu Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met Glu Ala | | | | | | |
| | | 435 | | 440 | | 445 |
| Thr Gly Val Arg Arg Asp Val Ala Tyr Leu Gln Ala Leu Ser Leu Glu | | | | | | |
| | | 450 | | 455 | | 460 |
| Leu Ala Glu Glu Ile Arg Arg Leu Glu Glu Glu Val Phe Arg Leu Ala | | | | | | |
| | | 465 | | 470 | | 475 |
| Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu | | | | | | |
| | | 485 | | 490 | | 495 |
| Phe Asp Glu Leu Arg Leu Pro Ala Leu Lys Lys Thr Lys Lys Thr Gly | | | | | | |
| | | 500 | | 505 | | 510 |
| Lys Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His | | | | | | |
| | | 515 | | 520 | | 525 |
| Pro Ile Val Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys Leu Lys | | | | | | |

<400> 549
atgggtgcgg atattggtga cctctttgag aggggaagagg tcgagcttga gtactttctca 60
ggaaagaaaa ttgccgttga tgctttcaac acgctatacc agttcatctc gataataagg 120
cagcctgacg gtacgccgtt aaaggactca cagggcagaa tcacctctca cctttccgga 180
atcctataca gagtctccaa catggctcag gtgggaatca ggccggtggt tgtattcgac 240
ggagagccac cggagttcaa gaaggctgaa attgaggaga ggaaaaagag aagggtgag 300
gcagaggaga tgtggattgc ggctttgcag gcaggagata aggacgcgaa aaagtatgct 360
caggctgcag ggagggttga cgagtacatt gttgactccg caaagacgct ttttaagttac 420
atggggattc cctttgtcga tgccccgtct gaaggagagg cgcaggctgc ttacatggca 480
gcaaaaggcg atgtggagta cacaggaagc caggattacg attctctgct cttcggaagc 540
ccgagactcg ccagaaatct cgcaataacg ggaaaaagga agcttcccgg caaaaatgtc 600
tatgtggatg taaagccgga gataataatt ctggaaagca acctcaaaag gctggggttg 660
acgagggagc agctcatcga catagcgatt ctggctcgga cggacgcaa tgaggggtgtg 720
aagggtgtcg gcgtcaagaa ggctttgaac tacatcaaga cctacggaga tatttttcagg 780
gcactcaagg ctctgaaagt aaatattgac cacgtagagg agataaggaa tttcttcctg 840
aatcctcctg tgactgacga ctacagaata gagttcaggg agcctgactt tgagaaggcc 900
atcgagttcc tgtgagagga gcacgacttc agcagggaga gggtcgagaa ggccttgag 960
aagctcaaag ctctgaagtc aaccagggcc acgcttgaga ggtgggttc 1008

<210> 550
<211> 336
<212> PRT
<213> Artificial

<220>
<223> Synthetic

<400> 550

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Gly | Ala | Asp | Ile | Gly | Asp | Leu | Phe | Glu | Arg | Glu | Glu | Val | Glu | Leu |
| 1 | | | | 5 | | | | 10 | | | | | | 15 | |
| Glu | Tyr | Phe | Ser | Gly | Lys | Lys | Ile | Ala | Val | Asp | Ala | Phe | Asn | Thr | Leu |
| | | | 20 | | | | 25 | | | | | | 30 | | |
| Tyr | Gln | Phe | Ile | Ser | Ile | Ile | Arg | Gln | Pro | Asp | Gly | Thr | Pro | Leu | Lys |
| | | | 35 | | | | 40 | | | | | 45 | | | |
| Asp | Ser | Gln | Gly | Arg | Ile | Thr | Ser | His | Leu | Ser | Gly | Ile | Leu | Tyr | Arg |

| | | | | |
|-----------------|---|---------------------|-----|----|
| 50 | | 55 | | 60 |
| Val Ser Asn Met | Val Glu Val Gly Ile Arg Pro | Val Phe Val Phe Asp | | |
| 65 | 70 | 75 | 80 | |
| Gly Glu Pro Pro | Glu Phe Lys Lys Ala Glu Ile Glu Glu Arg Lys Lys | | | |
| | 85 | 90 | 95 | |
| Arg Arg Ala Glu | Ala Glu Glu Met Trp Ile Ala Ala Leu Gln Ala Gly | | | |
| | 100 | 105 | 110 | |
| Asp Lys Asp Ala | Lys Lys Tyr Ala Gln Ala Ala Gly Arg Val Asp Glu | | | |
| | 115 | 120 | 125 | |
| Tyr Ile Val Asp | Ser Ala Lys Thr Leu Leu Ser Tyr Met Gly Ile Pro | | | |
| | 130 | 135 | 140 | |
| Phe Val Asp Ala | Pro Ser Glu Gly Glu Ala Gln Ala Ala Tyr Met Ala | | | |
| | 145 | 150 | 155 | |
| Ala Lys Gly Asp | Val Glu Tyr Thr Gly Ser Gln Asp Tyr Asp Ser Leu | | | |
| | 165 | 170 | 175 | |
| Leu Phe Gly Ser | Pro Arg Leu Ala Arg Asn Leu Ala Ile Thr Gly Lys | | | |
| | 180 | 185 | 190 | |
| Arg Lys Leu Pro | Gly Lys Asn Val Tyr Val Asp Val Lys Pro Glu Ile | | | |
| | 195 | 200 | 205 | |
| Ile Ile Leu Glu | Ser Asn Leu Lys Arg Leu Gly Leu Thr Arg Glu Gln | | | |
| | 210 | 215 | 220 | |
| Leu Ile Asp Ile | Ala Ile Leu Val Gly Thr Asp Ala Asn Glu Gly Val | | | |
| | 225 | 230 | 235 | |
| Lys Gly Val Gly | Val Lys Lys Ala Leu Asn Tyr Ile Lys Thr Tyr Gly | | | |
| | 245 | 250 | 255 | |
| Asp Ile Phe Arg | Ala Leu Lys Ala Leu Lys Val Asn Ile Asp His Val | | | |
| | 260 | 265 | 270 | |
| Glu Glu Ile Arg | Asn Phe Phe Leu Asn Pro Pro Val Thr Asp Asp Tyr | | | |
| | 275 | 280 | 285 | |
| Arg Ile Glu Phe | Arg Glu Pro Asp Phe Glu Lys Ala Ile Glu Phe Leu | | | |
| | 290 | 295 | 300 | |
| Cys Glu Glu His | Asp Phe Ser Arg Glu Arg Val Glu Lys Ala Leu Glu | | | |
| | 305 | 310 | 315 | |
| Lys Leu Lys Ala | Leu Lys Ser Thr Gln Ala Thr Leu Glu Arg Trp Phe | | | |
| | 325 | 330 | 335 | |

<210> 551
 <211> 33
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic
 <400> 551
 ctgggtcggga cggacaggaa tgagggtgtg aag 33

<210> 552
 <211> 1008
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic
 <400> 552
 atgggtgcgg atattggtga cctctttgag agggaagagg tcgagcttga gtactttctca 60
 ggaaagaaaa ttgccgttga tgctttcaac acgctatacc agttcatctc gataataagg 120
 cagcctgacg gtacgccgtt aaaggactca cagggcagaa tcacctctca cctttccgga 180
 atcctataca gagtctccaa catggctcag gtgggaatca ggccggtgtt tgtattcgac 240
 ggagagccac cggagttcaa gaaggctgaa attgaggaga ggaaaaagag aagggtctgag 300
 gcagaggaga tgtggattgc ggctttgcag gcaggagata aggacgcgaa aaagtatgct 360
 caggctgcag ggagggttga cgagtacatt gttgactcgc caaagacgct ttttaagttac 420
 atggggattc cctttgtcga tgccccgtct gaaggagagg cgcaggctgc ttacatggca 480
 gcaaaaggcg atgtggagta cacaggaagc caggattacg attctctgct cttcgggaagc 540
 ccgagactcg ccagaaatct cgcaataacg ggaaaaagga agcttcccgc caaaaatgct 600
 tatgtggatg taaagccgga gataataatt ctggaaagca acctcaaaag gctggggttg 660
 acgagggagc agctcatcga catagcgatt ctggtcggga cggacaggaa tgagggtgtg 720
 aagggtgtcg gcgtaagaa ggctttgaac tacatcaaga cctacggaga tattttcagg 780
 gcaactcaagg ctctgaaagt aaatattgac cacgtagagg agataaggaa tttcttcctg 840
 aatcctcctg tgactgacga ctacagaata gagttcaggg agcctgactt tgagaaggcc 900
 atcgagttcc tgtgagagga gcacgacttc agcagggaga gggtcgagaa ggccttgagg 960
 aagctcaaag ctctgaagtc aaccaggcc acgcttgaga ggtgggttc 1008

<210> 553
 <211> 336
 <212> PRT
 <213> Artificial

<220>

<223> Synthetic

<400> 553

Met Gly Ala Asp Ile Gly Asp Leu Phe Glu Arg Glu Glu Val Glu Leu
1 5 10 15
Glu Tyr Phe Ser Gly Lys Lys Ile Ala Val Asp Ala Phe Asn Thr Leu
20 25 30
Tyr Gln Phe Ile Ser Ile Ile Arg Gln Pro Asp Gly Thr Pro Leu Lys
35 40 45
Asp Ser Gln Gly Arg Ile Thr Ser His Leu Ser Gly Ile Leu Tyr Arg
50 55 60
Val Ser Asn Met Val Glu Val Gly Ile Arg Pro Val Phe Val Phe Asp
65 70 75 80
Gly Glu Pro Pro Glu Phe Lys Lys Ala Glu Ile Glu Glu Arg Lys Lys
85 90 95
Arg Arg Ala Glu Ala Glu Glu Met Trp Ile Ala Ala Leu Gln Ala Gly
100 105 110
Asp Lys Asp Ala Lys Lys Tyr Ala Gln Ala Ala Gly Arg Val Asp Glu
115 120 125
Tyr Ile Val Asp Ser Ala Lys Thr Leu Leu Ser Tyr Met Gly Ile Pro
130 135 140
Phe Val Asp Ala Pro Ser Glu Gly Glu Ala Gln Ala Ala Tyr Met Ala
145 150 155 160
Ala Lys Gly Asp Val Glu Tyr Thr Gly Ser Gln Asp Tyr Asp Ser Leu
165 170 175
Leu Phe Gly Ser Pro Arg Leu Ala Arg Asn Leu Ala Ile Thr Gly Lys
180 185 190
Arg Lys Leu Pro Gly Lys Asn Val Tyr Val Asp Val Lys Pro Glu Ile
195 200 205
Ile Ile Leu Glu Ser Asn Leu Lys Arg Leu Gly Leu Thr Arg Glu Gln
210 215 220
Leu Ile Asp Ile Ala Ile Leu Val Gly Thr Asp Arg Asn Glu Gly Val
225 230 235 240
Lys Gly Val Gly Val Lys Lys Ala Leu Asn Tyr Ile Lys Thr Tyr Gly
245 250 255
Asp Ile Phe Arg Ala Leu Lys Ala Leu Lys Val Asn Ile Asp His Val
260 265 270
Glu Glu Ile Arg Asn Phe Phe Leu Asn Pro Pro Val Thr Asp Asp Tyr

| | | | |
|--|------------------------|-----|-----|
| 275 | 280 | 285 | |
| Arg Ile Glu Phe Arg Glu Pro Asp Phe Glu Lys Ala Ile Glu Phe Leu | | | |
| 290 | 295 | 300 | |
| Cys Glu Glu His Asp Phe Ser Arg Glu Arg Val Glu Lys Ala Leu Glu | | | |
| 305 | 310 | 315 | 320 |
| Lys Leu Lys Ala Leu Lys Ser Thr Gln Ala Thr Leu Glu Arg Trp Phe | | | |
| | 325 | 330 | 335 |
| <210> | 554 | | |
| <211> | 20 | | |
| <212> | DNA | | |
| <213> | Artificial | | |
| <220> | | | |
| <223> | Synthetic | | |
| <400> | 554 | | |
| gaaccacctc tcaagcgtgg | | | 20 |
| <210> | 555 | | |
| <211> | 36 | | |
| <212> | DNA | | |
| <213> | Artificial | | |
| <220> | | | |
| <223> | Synthetic | | |
| <400> | 555 | | |
| acgcttgaga ggtggttcct ggaggaggcc ccctgg | | | 36 |
| <210> | 556 | | |
| <211> | 1011 | | |
| <212> | DNA | | |
| <213> | Archaeoglobus fulgidus | | |
| <400> | 556 | | |
| atgggtgcgg atattggtga cctctttgag agggaagagg tcgagcttga gtactttctca | | | 60 |
| ggaaagaaaa ttgccgttga tgctttcaac acgctatacc agttcatctc gataataagg | | | 120 |
| cagcctgacg gtacgccgtt aaaggactca cagggcagaa tcacctctca cctttccgga | | | 180 |
| atcctataca gagtctccaa catgggtcgag gtgggaatca ggccggtggt tgtattcgac | | | 240 |
| ggagagccac cggagttcaa gaaggctgaa attgaggaga ggaaaaagag aagggtgag | | | 300 |
| gcagaggaga tgtggattgc ggctttgcag gcaggagata aggacgcgaa aaagtatgct | | | 360 |
| caggctgcag ggagggttga cgagtacatt gttgactccg caaagacgct ttttaagttac | | | 420 |
| atggggattc cctttgtcga tgccccgtct gaaggagagg cgcaggctgc ttacatggca | | | 480 |

gcaaaaggcg atgtggagta cacaggaagc caggattacg attctctgct cttcgggaagc 540
 ccgagactcg ccagaaatct cgcaataacg ggaaaaagga agcttcccgg caaaaatgtc 600
 tatgtggatg taaagccgga gataataatt ctggaaagca acctcaaaag gctgggtttg 660
 acgaggggagc agctcatcga catagcgatt ctggtcggga cggactacaa tgaggggtgtg 720
 aaggggtgtcg gcgtcaagaa ggctttgaac tacatcaaga cctacggaga tatttttcagg 780
 gcactcaagg ctctgaaagt aaatattgac cacgtagagg agataaggaa tttcttcctg 840
 aatcctcctg tgactgacga ctacagaata gagttcaggg agcctgactt tgagaaggcc 900
 atcgagttcc tgtgcgagga gcacgacttc agcagggaga gggtcgagaa ggccttggag 960
 aagctcaaag ctctgaagtc aaccaggcc acgcttgaga ggtgggttctg a 1011

<210> 557
 <211> 17
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 557
 taatctgtat caggctg 17

<210> 558
 <211> 2526
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 558
 atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctgggtggac 60
 ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
 gaaccggtgc aggcggtcta cggtctcgcc aagagcctcc tcaaggccct gaaggaggac 180
 ggggtacaagg ccgtcttcgt ggtctttgac gccaaaggccc cctccttcctg ccacgaggcc 240
 tacgaggcct acaaggcggg gagggccccc acccccagag acttcccccg gcagctcgcc 300
 ctcatcaagg agctggtgga cctcctgggg ttaccgccg tcgaggtccc cggctacgag 360
 gcggacgacg ttctcgccac cctggccaag aaggcggaaa aggaggggta cgaggtgctc 420
 atcctcaccg ccgaccgcga cctctaccaa ctgctctccg accgcgtcgc cgtcctccac 480
 cccgagggcc acctcatcac cccggagtgg ctttggggaga agtacggcct caggccggag 540

| | | | | | | |
|------------|-------------|-------------|-------------|-------------|-------------|------|
| cagtgggtgg | acttccgcgc | cctcgtgggg | gacccctccg | acaacctccc | cggggtcaag | 600 |
| ggcatcgggg | agaagaccgc | cctcaagctc | ctcaaggagt | ggggaagcct | ggaaaacctc | 660 |
| ctcaagaacc | tggaccgggt | aaagccagaa | aacgtccggg | agaagatcaa | ggccccacctg | 720 |
| gaagacctca | ggctctcctt | ggagctctcc | cggggtgcga | ccgacctccc | cctggagggtg | 780 |
| gacctcgccc | aggggctggg | gcccgaaccg | gaggggctta | gggccttcct | ggagagggtg | 840 |
| gagttcgga | gcctcctcca | cgagttcggc | ctcctggagg | cccccgcccc | cctggaggag | 900 |
| gccccctggc | ccccgccgga | aggggccttc | gtgggcttcg | tcctctcccg | ccccgagccc | 960 |
| atgtgggcgg | agcttaaagc | cctggccgcc | tgcaggggcg | gccgcgtgca | ccgggcagca | 1020 |
| gaccccttgg | cggggctaaa | ggacctcaag | gaggtccggg | gcctcctcgc | caaggacctc | 1080 |
| gccgtcttgg | cctcgaggga | ggggctagac | ctcgtgcccc | gggacgaccc | catgctcttc | 1140 |
| gcctacctcc | tggaccttc | gaacaccacc | cccagggggg | tggcgcggcg | ctacggggggg | 1200 |
| gagtggacgg | aggacgccgc | ccaccggggc | ctcctctcgg | agaggctcca | tcggaacctc | 1260 |
| cttaagcgcc | tcgaggggga | ggagaagctc | ctttggctct | accacgaggt | ggaaaagccc | 1320 |
| ctctcccggg | tcctggccca | tatggaggcc | accgggggtac | ggcgggacgt | ggcctacctt | 1380 |
| caggcccttt | ccttgagct | tgcggaggag | atccgccgcc | tcgaggagga | ggtcttcgcg | 1440 |
| ttggcgggcc | accccttcaa | cctcaactcc | cgggaccagc | tggaaaagggt | gctctttgac | 1500 |
| gagcttaggc | ttcccgctt | gaagaagacg | aagaagacag | gcaagcgctc | caccagcgcc | 1560 |
| gcggtgctgg | agggcctacg | ggaggccccc | cccatcgtag | agaagatcct | ccagcaccgg | 1620 |
| gagctcacca | agctcaagaa | cacctacgtg | gacccctccc | caagcctcgt | ccacccgagg | 1680 |
| acgggcccgc | tccacacccg | cttcaaccag | acggccacgg | ccacggggag | gcttagtagc | 1740 |
| tccgacccca | acctgcagaa | catccccgtc | cgcacccctt | tgggccagag | gatccgcccg | 1800 |
| gccttcgtgg | ccgaggcggg | ttgggcgttg | gtggccctgg | actatagcca | gatagagctc | 1860 |
| cgcgctctcg | cccacctctc | cggggacgaa | aacctgatca | gggtcttcca | ggaggggaag | 1920 |
| gacatccaca | cccagaccgc | aagctggatg | ttcggcgctc | ccccggaggc | cgtggacccc | 1980 |
| ctgatgcgcc | gggcggccaa | gacgggtgaac | ttcggcgctc | tctacggcat | gtccgcccac | 2040 |
| aggctctccc | aggagcttgc | catccccctac | gaggaggcgg | tggcctttat | agagcgctac | 2100 |
| ttccaaaget | tccccaagggt | gcgggccttg | atagaaaaga | ccctggagga | ggggaggaag | 2160 |
| cggggctacg | tggaaaccct | cttcggaaga | aggcgctacg | tgcccgacct | caacgcccgg | 2220 |

| | | | | | | |
|-------------|------------|------------|------------|------------|------------|------|
| tgggcgagc | ttaaagccct | ggccgcctgc | aggggcggcc | gcgtgcaccg | ggcagcagac | 1140 |
| cccttggcgg | ggctaaagga | cctcaaggag | gtccggggcc | tcctcgccaa | ggacctcgcc | 1200 |
| gtcttggcct | cgagggaggg | gctagacctc | gtgccggggg | acgaccccat | gctcctcgcc | 1260 |
| tacctctgg | acccttcgaa | caccaccccc | gagggggtgg | cgcggcgcta | cgggggggag | 1320 |
| tggacggagg | acgccgccca | ccgggccttc | ctctcggaga | ggctccatcg | gaacctcctt | 1380 |
| aagcgctcg | agggggagga | gaagctcctt | tggctctacc | acgaggtgga | aaagcccctc | 1440 |
| tcccgggtcc | tggcccatat | ggaggccacc | ggggtacggc | gggacgtggc | ctaccttcag | 1500 |
| gccctttccc | tggagcttgc | ggaggagatc | cgccgcctcg | aggaggaggt | cttcgccttg | 1560 |
| gcggggccacc | ccttcaacct | caactcccgg | gaccagctgg | aaagggtgct | ctttgacgag | 1620 |
| cttaggcttc | ccgccttgaa | gaagacgaag | aagacaggca | agcgctccac | cagcgccgcy | 1680 |
| gtgctggagg | ccctacggga | ggcccacccc | atcgtggaga | agatcctcca | gcaccgggag | 1740 |
| ctcaccaagc | tcaagaacac | ctacgtggac | cccctcccaa | gcctcgtoea | cccaggagcy | 1800 |
| ggcgcgctcc | acaccgctt | caaccagacg | gccacggcca | cggggagggt | tagtagctcc | 1860 |
| gaccccaacc | tgcagaacat | cccgtccgc | accccttgg | gccagaggat | ccgccgggcc | 1920 |
| ttcgtggccg | aggcggttg | ggcgttggtg | gccctggact | atagccagat | agagctccgc | 1980 |
| gtcctcgccc | acctctccgg | ggacgaaaac | ctgatcaggg | tcttccagga | ggggaaggac | 2040 |
| atccacaccc | agaccgcaag | ctggatgttc | ggcgtccccc | cggaggccgt | ggacccctg | 2100 |
| atgcgcgggg | cggccaagac | ggtgaacttc | ggcgtcctct | acggcatgtc | cgcccatagg | 2160 |
| ctctcccagg | agcttgccat | cccctacgag | gaggcggtgg | cctttataga | gcgctacttc | 2220 |
| caaagcttcc | ccaaggtgcy | ggcctggata | gaaaagacct | tggaggaggg | gaggaagcgg | 2280 |
| ggctacgtgg | aaacctcctt | cggaagaagg | cgctacgtgc | ccgacctcaa | cgcccggtg | 2340 |
| aagagcgtca | gggaggccgc | ggagcgcacg | gccttcaaca | tgcccgtcca | gggcaccgcc | 2400 |
| gccgacctca | tgaagctcgc | catggtgaag | ctcttcccc | gcctccggga | gatggggggc | 2460 |
| cgcatgctcc | tccaggtcgc | caacgagctc | ctcctggagg | ccccccaagc | gcgggcccag | 2520 |
| gaggtggcgg | ctttggccaa | ggaggccatg | gagaaggcct | atcccctcgc | cgtgcccctg | 2580 |
| gaggtggagg | tggggatggg | ggaggactgg | ctttccgcca | agggtcacca | ccaccaccac | 2640 |
| cac | | | | | | 2643 |

<210> 560

<211> 881
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 560

```

Met Gly Ala Asp Ile Gly Asp Leu Phe Glu Arg Glu Glu Val Glu Leu
1           5           10           15

Glu Tyr Phe Ser Gly Lys Lys Ile Ala Val Asp Ala Phe Asn Thr Leu
          20           25           30

Tyr Gln Phe Ile Ser Ile Ile Arg Gln Pro Asp Gly Thr Pro Leu Lys
          35           40           45

Asp Ser Gln Gly Arg Ile Thr Ser His Leu Ser Gly Ile Leu Tyr Arg
          50           55           60

Val Ser Asn Met Val Glu Val Gly Ile Arg Pro Val Phe Val Phe Asp
65           70           75           80

Gly Glu Pro Pro Glu Phe Lys Lys Ala Glu Ile Glu Glu Arg Lys Lys
          85           90           95

Arg Arg Ala Glu Ala Glu Glu Met Trp Ile Ala Ala Leu Gln Ala Gly
          100          105          110

Asp Lys Asp Ala Lys Lys Tyr Ala Gln Ala Ala Gly Arg Val Asp Glu
          115          120          125

Tyr Ile Val Asp Ser Ala Lys Thr Leu Leu Ser Tyr Met Gly Ile Pro
          130          135          140

Phe Val Asp Ala Pro Ser Glu Gly Glu Ala Gln Ala Ala Tyr Met Ala
145          150          155          160

Ala Lys Gly Asp Val Glu Tyr Thr Gly Ser Gln Asp Tyr Asp Ser Leu
          165          170          175

Leu Phe Gly Ser Pro Arg Leu Ala Arg Asn Leu Ala Ile Thr Gly Lys
          180          185          190

Arg Lys Leu Pro Gly Lys Asn Val Tyr Val Asp Val Lys Pro Glu Ile
          195          200          205

Ile Ile Leu Glu Ser Asn Leu Lys Arg Leu Gly Leu Thr Arg Glu Gln
          210          215          220

Leu Ile Asp Ile Ala Ile Leu Val Gly Thr Asp Tyr Asn Glu Gly Val
225          230          235          240

Lys Gly Val Gly Val Lys Lys Ala Leu Asn Tyr Ile Lys Thr Tyr Gly
          245          250          255

```

10259-9999-0340

| | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| Val | Leu | Glu | Ala | Leu | Arg | Glu | Ala | His | Pro | Ile | Val | Glu | Lys | Ile | Leu | |
| | | | | 565 | | | | | 570 | | | | | 575 | | |
| Gln | His | Arg | Glu | Leu | Thr | Lys | Leu | Lys | Asn | Thr | Tyr | Val | Asp | Pro | Leu | |
| | | | 580 | | | | | 585 | | | | | 590 | | | |
| Pro | Ser | Leu | Val | His | Pro | Arg | Thr | Gly | Arg | Leu | His | Thr | Arg | Phe | Asn | |
| | | 595 | | | | | 600 | | | | | 605 | | | | |
| Gln | Thr | Ala | Thr | Ala | Thr | Gly | Arg | Leu | Ser | Ser | Ser | Asp | Pro | Asn | Leu | |
| | 610 | | | | | 615 | | | | | 620 | | | | | |
| Gln | Asn | Ile | Pro | Val | Arg | Thr | Pro | Leu | Gly | Gln | Arg | Ile | Arg | Arg | Ala | |
| 625 | | | | | 630 | | | | | 635 | | | | | 640 | |
| Phe | Val | Ala | Glu | Ala | Gly | Trp | Ala | Leu | Val | Ala | Leu | Asp | Tyr | Ser | Gln | |
| | | | | 645 | | | | | 650 | | | | | 655 | | |
| Ile | Glu | Leu | Arg | Val | Leu | Ala | His | Leu | Ser | Gly | Asp | Glu | Asn | Leu | Ile | |
| | | | 660 | | | | | 665 | | | | | 670 | | | |
| Arg | Val | Phe | Gln | Glu | Gly | Lys | Asp | Ile | His | Thr | Gln | Thr | Ala | Ser | Trp | |
| | | 675 | | | | | 680 | | | | | 685 | | | | |
| Met | Phe | Gly | Val | Pro | Pro | Glu | Ala | Val | Asp | Pro | Leu | Met | Arg | Arg | Ala | |
| | 690 | | | | | 695 | | | | | 700 | | | | | |
| Ala | Lys | Thr | Val | Asn | Phe | Gly | Val | Leu | Tyr | Gly | Met | Ser | Ala | His | Arg | |
| 705 | | | | 710 | | | | | | 715 | | | | | 720 | |
| Leu | Ser | Gln | Glu | Leu | Ala | Ile | Pro | Tyr | Glu | Glu | Ala | Val | Ala | Phe | Ile | |
| | | | 725 | | | | | | 730 | | | | | 735 | | |
| Glu | Arg | Tyr | Phe | Gln | Ser | Phe | Pro | Lys | Val | Arg | Ala | Trp | Ile | Glu | Lys | |
| | | 740 | | | | | | 745 | | | | | 750 | | | |
| Thr | Leu | Glu | Glu | Gly | Arg | Lys | Arg | Gly | Tyr | Val | Glu | Thr | Leu | Phe | Gly | |
| | 755 | | | | | | 760 | | | | | 765 | | | | |
| Arg | Arg | Arg | Tyr | Val | Pro | Asp | Leu | Asn | Ala | Arg | Val | Lys | Ser | Val | Arg | |
| | 770 | | | | | 775 | | | | | 780 | | | | | |
| Glu | Ala | Ala | Glu | Arg | Met | Ala | Phe | Asn | Met | Pro | Val | Gln | Gly | Thr | Ala | |
| 785 | | | | | 790 | | | | | 795 | | | | | 800 | |
| Ala | Asp | Leu | Met | Lys | Leu | Ala | Met | Val | Lys | Leu | Phe | Pro | Arg | Leu | Arg | |
| | | | 805 | | | | | | 810 | | | | | 815 | | |
| Glu | Met | Gly | Ala | Arg | Met | Leu | Leu | Gln | Val | Ala | Asn | Glu | Leu | Leu | Leu | |
| | | 820 | | | | | | 825 | | | | 830 | | | | |
| Glu | Ala | Pro | Gln | Ala | Arg | Ala | Glu | Glu | Val | Ala | Ala | Leu | Ala | Lys | Glu | |
| | | 835 | | | | | 840 | | | | | 845 | | | | |
| Ala | Met | Glu | Lys | Ala | Tyr | Pro | Leu | Ala | Val | Pro | Leu | Glu | Val | Glu | Val | |
| | 850 | | | | | 855 | | | | | 860 | | | | | |

Gly Met Gly Glu Asp Trp Leu Ser Ala Lys Gly His His His His His
 865 870 875 880

His

<210> 561
 <211> 21
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 561
 gggtgacttc agagctttga g 21

<210> 562
 <211> 36
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 562
 aaagctctga agtcaaccct ggaggaggcc ccctgg 36

<210> 563
 <211> 2619
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 563
 atgggtgcgg atattggtga cctctttgag agggaagagg tcgagcttga gtactttctca 60
 ggaaagaaaa ttgccgttga tgctttcaac acgctatacc agttcatctc gataataagg 120
 cagcctgacg gtacgccgtt aaaggactca cagggcagaa tcacctctca cctttccgga 180
 atcctataca gagtctccaa catggctgag gtgggaatca ggccgggtgtt tgtattcgac 240
 ggagagccac cggagttcaa gaaggctgaa attgaggaga ggaaaaagag aagggtgag 300
 gcagaggaga tgtggattgc ggctttgcag gcaggagata aggacgcgaa aaagtatgct 360
 caggctgcag ggagggttga cgagtacatt gttgactccg caaagacgct ttttaagttac 420
 atggggattc cctttgtcga tgccccgtct gaaggagagg cgcaggctgc ttacatggca 480
 gcaaaaggcg atgtggagta cacaggaagc caggattacg attctctgct cttcgaagc 540

| | |
|---|------|
| ccgagactcg ccagaaatct cgcaataacg ggaaaaagga agcttcccgg caaaaatgtc | 600 |
| tatgtggatg taaagccgga gataataatt ctggaaagca acctcaaaag gctgggtttg | 660 |
| acgagggagc agctcatcga catagcgatt ctggtcggga cggactacaa tgaggggtgtg | 720 |
| aaggggtgtcg gcgtcaagaa ggctttgaac tacatcaaga cctacggaga tatttttcagg | 780 |
| gcactcaagg ctctgaaagt aaatattgac cacgtagagg agataaggaa tttcttctcg | 840 |
| aatcctcctg tgactgacga ctacagaata gagttcaggg agcctgactt tgagaaggcc | 900 |
| atcgagttcc tgtgcgagga gcacgacttc agcagggaga gggtcgagaa ggccttgag | 960 |
| aagctcaaag ctctgaagtc aaccctggag gaggccccct ggcccccgcc ggaaggggccc | 1020 |
| ttcgtgggct tcgtcctctc ccgccccgag cccatgtggg cggagcttaa agccctggcc | 1080 |
| gcctgcaggg gcggcccggt gcaccgggca gcagaccctc tggcggggct aaaggacctc | 1140 |
| aaggaggtcc ggggcctcct cgccaaggac ctgcgcgtct tggcctcgag ggaggggcta | 1200 |
| gacctcgtgc ccggggacga ccccatgctc ctgcctacc tcctggacct ttcgaacacc | 1260 |
| acccccgagg ggggtggcgcg gcgctacggg ggggagtgga cggaggacgc cgcccaccgg | 1320 |
| gccctcctct cggagaggct ccatcggaac ctcttaagc gcctcgaggg ggaggagaag | 1380 |
| ctcctttggc tctaccacga ggtggaaaag cccctctccc gggtcctggc ccatatggag | 1440 |
| gccaccgggg tacggcgagg cgtggcctac cttcaggccc tttccctgga gcttgcgag | 1500 |
| gagatccgcc gcctcgagga ggaggtcttc cgcttggcgg gccaccctt caacctcaac | 1560 |
| tcccgggacc agctggaaaag ggtgctcttt gacgagctta ggcttcccgc cttgaagaag | 1620 |
| acgaagaaga caggcaagcg ctccaccagc gccgcgggtgc tggaggccct acgggaggcc | 1680 |
| caccccatcg tggagaagat cctccagcac cgggagctca ccaagctcaa gaacacctac | 1740 |
| gtggaccccc tccaagcct cgtccaccgc aggacgggcc gcctccacac ccgcttcaac | 1800 |
| cagacggcca cggccacggg gaggcttagt agctccgacc ccaacctgca gaacatcccc | 1860 |
| gtccgcaccc ccttgggcca gaggatccgc cgggccttcg tggccgaggc gggttgggcg | 1920 |
| ttggtggccc tggactatag ccagatagag ctccgcgtcc tcgcccacct ctccggggac | 1980 |
| gaaaacctga tcagggtctt ccaggagggg aaggacatcc acaccagac cgcaagctgg | 2040 |
| atgttcggcg tcccccgga ggccgtggac cccctgatgc gccgggcggc caagacgggtg | 2100 |
| aacttcggcg tcctctacgg catgtccgcc cataggctct cccaggagct tgccatcccc | 2160 |
| tacgaggagg cgggtggcctt tatagagcgc tacttccaaa gcttccccaa ggtgcggggc | 2220 |
| tggatagaaa agaccctgga ggaggggagg aagcggggct acgtggaaac cctcttcgga | 2280 |

agaaggcgct acgtgcccga cctcaacgcc cgggtgaaga gcgtcagggg ggccgcggag 2340
 cgcatggcct tcaacatgcc cgtccagggc accgcgcgcg acctcatgaa gctcgccatg 2400
 gtgaagctct tccccgcct ccgggagatg ggggcccgcg tgctcctcca ggtcgccaac 2460
 gagctcctcc tggaggcccc ccaagcgcgg gccgaggagg tggcggcttt ggccaaggag 2520
 gccatggaga aggcctatcc cctcgccgtg cccctggagg tggaggtggg gatgggggag 2580
 gactggcttt ccgccaaggg tcaccaccac caccaccac 2619

<210> 564
 <211> 873
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 564

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Gly | Ala | Asp | Ile | Gly | Asp | Leu | Phe | Glu | Arg | Glu | Glu | Val | Glu | Leu |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Glu | Tyr | Phe | Ser | Gly | Lys | Lys | Ile | Ala | Val | Asp | Ala | Phe | Asn | Thr | Leu |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Tyr | Gln | Phe | Ile | Ser | Ile | Ile | Arg | Gln | Pro | Asp | Gly | Thr | Pro | Leu | Lys |
| | | | 35 | | | | 40 | | | | | 45 | | | |
| Asp | Ser | Gln | Gly | Arg | Ile | Thr | Ser | His | Leu | Ser | Gly | Ile | Leu | Tyr | Arg |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| Val | Ser | Asn | Met | Val | Glu | Val | Gly | Ile | Arg | Pro | Val | Phe | Val | Phe | Asp |
| 65 | | | | | 70 | | | | | 75 | | | | | 80 |
| Gly | Glu | Pro | Pro | Glu | Phe | Lys | Lys | Ala | Glu | Ile | Glu | Glu | Arg | Lys | Lys |
| | | | | 85 | | | | | 90 | | | | | 95 | |
| Arg | Arg | Ala | Glu | Ala | Glu | Glu | Met | Trp | Ile | Ala | Ala | Leu | Gln | Ala | Gly |
| | | | 100 | | | | | 105 | | | | | 110 | | |
| Asp | Lys | Asp | Ala | Lys | Lys | Tyr | Ala | Gln | Ala | Ala | Gly | Arg | Val | Asp | Glu |
| | | 115 | | | | | 120 | | | | | 125 | | | |
| Tyr | Ile | Val | Asp | Ser | Ala | Lys | Thr | Leu | Leu | Ser | Tyr | Met | Gly | Ile | Pro |
| | 130 | | | | | 135 | | | | | | 140 | | | |
| Phe | Val | Asp | Ala | Pro | Ser | Glu | Gly | Glu | Ala | Gln | Ala | Ala | Tyr | Met | Ala |
| 145 | | | | | 150 | | | | | 155 | | | | | 160 |
| Ala | Lys | Gly | Asp | Val | Glu | Tyr | Thr | Gly | Ser | Gln | Asp | Tyr | Asp | Ser | Leu |
| | | | | 165 | | | | | 170 | | | | | 175 | |

| | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Leu | Phe | Gly | Ser | Pro | Arg | Leu | Ala | Arg | Asn | Leu | Ala | Ile | Thr | Gly | Lys | 180 | 185 | 190 |
| Arg | Lys | Leu | Pro | Gly | Lys | Asn | Val | Tyr | Val | Asp | Val | Lys | Pro | Glu | Ile | 195 | 200 | 205 |
| Ile | Ile | Leu | Glu | Ser | Asn | Leu | Lys | Arg | Leu | Gly | Leu | Thr | Arg | Glu | Gln | 210 | 215 | 220 |
| Leu | Ile | Asp | Ile | Ala | Ile | Leu | Val | Gly | Thr | Asp | Tyr | Asn | Glu | Gly | Val | 225 | 230 | 235 |
| Lys | Gly | Val | Gly | Val | Lys | Lys | Ala | Leu | Asn | Tyr | Ile | Lys | Thr | Tyr | Gly | 245 | 250 | 255 |
| Asp | Ile | Phe | Arg | Ala | Leu | Lys | Ala | Leu | Lys | Val | Asn | Ile | Asp | His | Val | 260 | 265 | 270 |
| Glu | Glu | Ile | Arg | Asn | Phe | Phe | Leu | Asn | Pro | Pro | Val | Thr | Asp | Asp | Tyr | 275 | 280 | 285 |
| Arg | Ile | Glu | Phe | Arg | Glu | Pro | Asp | Phe | Glu | Lys | Ala | Ile | Glu | Phe | Leu | 290 | 295 | 300 |
| Cys | Glu | Glu | His | Asp | Phe | Ser | Arg | Glu | Arg | Val | Glu | Lys | Ala | Leu | Glu | 305 | 310 | 315 |
| Lys | Leu | Lys | Ala | Leu | Lys | Ser | Thr | Leu | Glu | Glu | Ala | Pro | Trp | Pro | Pro | 325 | 330 | 335 |
| Pro | Glu | Gly | Ala | Phe | Val | Gly | Phe | Val | Leu | Ser | Arg | Pro | Glu | Pro | Met | 340 | 345 | 350 |
| Trp | Ala | Glu | Leu | Lys | Ala | Leu | Ala | Ala | Cys | Arg | Gly | Gly | Arg | Val | His | 355 | 360 | 365 |
| Arg | Ala | Ala | Asp | Pro | Leu | Ala | Gly | Leu | Lys | Asp | Leu | Lys | Glu | Val | Arg | 370 | 375 | 380 |
| Gly | Leu | Leu | Ala | Lys | Asp | Leu | Ala | Val | Leu | Ala | Ser | Arg | Glu | Gly | Leu | 385 | 390 | 395 |
| Asp | Leu | Val | Pro | Gly | Asp | Asp | Pro | Met | Leu | Leu | Ala | Tyr | Leu | Leu | Asp | 405 | 410 | 415 |
| Pro | Ser | Asn | Thr | Thr | Pro | Glu | Gly | Val | Ala | Arg | Arg | Tyr | Gly | Gly | Glu | 420 | 425 | 430 |
| Trp | Thr | Glu | Asp | Ala | Ala | His | Arg | Ala | Leu | Leu | Ser | Glu | Arg | Leu | His | 435 | 440 | 445 |
| Arg | Asn | Leu | Leu | Lys | Arg | Leu | Glu | Gly | Glu | Glu | Lys | Leu | Leu | Trp | Leu | 450 | 455 | 460 |
| Tyr | His | Glu | Val | Glu | Lys | Pro | Leu | Ser | Arg | Val | Leu | Ala | His | Met | Glu | 465 | 470 | 475 |

Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met
785 790 795 800

Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu Leu
805 810 815

Gln Val Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala Glu
820 825 830

Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro Leu
835 840 845

Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser
850 855 860

Ala Lys Gly His His His His His His
865 870

<210> 565
<211> 2643
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 565
atgggtgcgg atattggtga cctctttgag aggaagagg tcgagcttga gtactttctca 60
ggaaagaaaa ttgccgttga tgctttcaac acgctatacc agttcatctc gataataagg 120
cagcctgacg gtacgccgtt aaaggactca cagggcagaa tcacctctca cttttccgga 180
atcctataca gagtctccaa catggtcgag gtgggaatca ggccggtgtt tgtattcgac 240
ggagagccac cggagttcaa gaaggctgaa attgaggaga ggaaaaagag aagggtcgag 300
gcagaggaga tgtggattgc ggctttgcag gcaggagata aggacgcgaa aaagtatgct 360
caggctgcag ggaggggttga cgagtacatt gttgactccg caaagacgct ttttaagttac 420
atggggattc cctttgtcga tgccccgtct gaaggagagg cgcaggctgc ttacatggca 480
gcaaaaggcg atgtggagta cacaggaagc caggattacg attctctgct cttcggaagc 540
ccgagactcg ccagaaatct cgcaataacg ggaaaaagga agcttcccg caaaaatgtc 600
tatgtggatg taaagccgga gataataatt ctggaaagca acctcaaaag gctgggtttg 660
acgagggagc agctcatcga catagcgatt ctggtcggga cggactacaa tgaggggtgtg 720
aagggtgtcg gcgtcaagaa ggctttgaac tacatcaaga cctacggaga tatttttcagg 780
gcactcaagg ctctgaaagt aaatattgac cacgtagagg agataaggaa tttcttcctg 840
aatcctcctg tgactgacga ctacagaata gagttcaggg agcctgactt tgagaaggcc 900

| | |
|---|------|
| atcgagttcc tgtgcgagga gcacgacttc agcagggaga gggtcgagaa ggccttggag | 960 |
| aagctcaaag ctctgaagtc aaccaggcc acgcttgaga ggtggttcct ggaggaggcc | 1020 |
| ccctggcccc cgccggaagg ggccttcgtg ggcttcgtcc tctcccgccc cgagcccatg | 1080 |
| tgggcgagc ttaaagccct ggccgcctgc agggcgggcc gcgtccaccg ggccccgag | 1140 |
| ccttataaag ccctcagga cctgaaggag gcgcgggggc ttctcgccaa agacctgagc | 1200 |
| gttctggccc tgagggaagg ccttggcctc ccgcccggcg acgaccccat gctcctcgcc | 1260 |
| tacctctgg acccttcgaa caccacccc gagggggtgg ccggcgcta cggcggggag | 1320 |
| tggacggagg aggcggggga gcgggcccgc cttccgaga ggctcttcgc caacctgctt | 1380 |
| aagaggcttg agggggagga gaggtcctt tggctttacc gggagggtga gagggccctt | 1440 |
| tccgctgtcc tggcccatat ggaggccacg ggggtgccc tggacgtggc ctatctcagg | 1500 |
| gccttgccc tggagggtggc cgaggagatc gccgcctcg aggcgaggt cttccgcctg | 1560 |
| gccggccacc ccttcaacct caactcccgg gaccagctgg aaagggtcct ctttgacgag | 1620 |
| ctagggttc ccgccatcaa gaagacgcaa aagaccggca agcgctccac cagcgccgc | 1680 |
| gtcctggagg ccctccgca gggccacccc atcgtggaga agatcctgca gtaccgggag | 1740 |
| ctaccaagc tgaagagcac ctacattgac cccttgccgg acctcatcca cccaggacg | 1800 |
| ggcgcctcc acaccgctt caaccagacg gccacggcca cgggcaggct aagtagctcc | 1860 |
| gatcccaacc tccagaacat ccccgctcgc acccgcttg ggcagaggat ccgcccggcc | 1920 |
| ttcatcgccg aggaggggtg gctattggtg gccctggact atagccagat agagctcagg | 1980 |
| gtgctggccc acctctccgg cgacgagaac ctgatccggg tcttcagga gggcgggac | 2040 |
| atccacacgg agaccgccag ctggatgttc ggcgtccccc gggaggccgt ggacccctg | 2100 |
| atgcgccggg cgccaagac catcaacttc ggggtcctct acggcatgtc ggcccaccgc | 2160 |
| ctctcccagg agctagccat cccttacgag gaggcccagg ccttcattga gcgctacttt | 2220 |
| cagagcttcc ccaaggtgcg ggctggatt gagaagacc tggaggaggg caggaggcgg | 2280 |
| gggtacgtgg agaccctctt cgcccgccgc cgctacgtgc cagacctaga ggccgggtg | 2340 |
| aagagcgtgc gggaggcggc cgagcgcatg gccttcaaca tgcccgcca gggcaccgcc | 2400 |
| gccgacctca tgaagctggc tatggtgaag ctcttcccca ggctggagga aatggggggc | 2460 |
| aggatgctcc ttcaggtcgc caacgagctg gtcctcgagg ccccaaaaga gagggcgag | 2520 |
| gccgtggccc ggctggccaa ggaggcatg gagggggtgt atccctggc cgtgcccctg | 2580 |
| gaggtggagg tggggatagg ggaggactgg ctctccgcca aggagcacca ccaccaccac | 2640 |

cac

2643

<210> 566
<211> 881
<212> PRT
<213> Artificial

<220>
<223> Synthetic

<400> 566

Met Gly Ala Asp Ile Gly Asp Leu Phe Glu Arg Glu Glu Val Glu Leu
1 5 10 15
Glu Tyr Phe Ser Gly Lys Lys Ile Ala Val Asp Ala Phe Asn Thr Leu
20 25 30
Tyr Gln Phe Ile Ser Ile Ile Arg Gln Pro Asp Gly Thr Pro Leu Lys
35 40 45
Asp Ser Gln Gly Arg Ile Thr Ser His Leu Ser Gly Ile Leu Tyr Arg
50 55 60
Val Ser Asn Met Val Glu Val Gly Ile Arg Pro Val Phe Val Phe Asp
65 70 75 80
Gly Glu Pro Pro Glu Phe Lys Lys Ala Glu Ile Glu Glu Arg Lys Lys
85 90 95
Arg Arg Ala Glu Ala Glu Glu Met Trp Ile Ala Ala Leu Gln Ala Gly
100 105 110
Asp Lys Asp Ala Lys Lys Tyr Ala Gln Ala Ala Gly Arg Val Asp Glu
115 120 125
Tyr Ile Val Asp Ser Ala Lys Thr Leu Leu Ser Tyr Met Gly Ile Pro
130 135 140
Phe Val Asp Ala Pro Ser Glu Gly Glu Ala Gln Ala Ala Tyr Met Ala
145 150 155 160
Ala Lys Gly Asp Val Glu Tyr Thr Gly Ser Gln Asp Tyr Asp Ser Leu
165 170 175
Leu Phe Gly Ser Pro Arg Leu Ala Arg Asn Leu Ala Ile Thr Gly Lys
180 185 190
Arg Lys Leu Pro Gly Lys Asn Val Tyr Val Asp Val Lys Pro Glu Ile
195 200 205
Ile Ile Leu Glu Ser Asn Leu Lys Arg Leu Gly Leu Thr Arg Glu Gln
210 215 220
Leu Ile Asp Ile Ala Ile Leu Val Gly Thr Asp Tyr Asn Glu Gly Val

| | | | | | | |
|---|--|-----|--|-----|--|-----|
| 225 | | 230 | | 235 | | 240 |
| Lys Gly Val Gly Val Lys Lys Ala Leu Asn Tyr Ile Lys Thr Tyr Gly | | | | | | |
| | | 245 | | 250 | | 255 |
| Asp Ile Phe Arg Ala Leu Lys Ala Leu Lys Val Asn Ile Asp His Val | | | | | | |
| | | 260 | | 265 | | 270 |
| Glu Glu Ile Arg Asn Phe Phe Leu Asn Pro Pro Val Thr Asp Asp Tyr | | | | | | |
| | | 275 | | 280 | | 285 |
| Arg Ile Glu Phe Arg Glu Pro Asp Phe Glu Lys Ala Ile Glu Phe Leu | | | | | | |
| | | 290 | | 295 | | 300 |
| Cys Glu Glu His Asp Phe Ser Arg Glu Arg Val Glu Lys Ala Leu Glu | | | | | | |
| 305 | | 310 | | 315 | | 320 |
| Lys Leu Lys Ala Leu Lys Ser Thr Gln Ala Thr Leu Glu Arg Trp Phe | | | | | | |
| | | 325 | | 330 | | 335 |
| Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu Gly Ala Phe Val Gly Phe | | | | | | |
| | | 340 | | 345 | | 350 |
| Val Leu Ser Arg Pro Glu Pro Met Trp Ala Glu Leu Lys Ala Leu Ala | | | | | | |
| | | 355 | | 360 | | 365 |
| Ala Cys Arg Gly Gly Arg Val His Arg Ala Pro Glu Pro Tyr Lys Ala | | | | | | |
| | | 370 | | 375 | | 380 |
| Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu Leu Ala Lys Asp Leu Ser | | | | | | |
| 385 | | 390 | | 395 | | 400 |
| Val Leu Ala Leu Arg Glu Gly Leu Gly Leu Pro Pro Gly Asp Asp Pro | | | | | | |
| | | 405 | | 410 | | 415 |
| Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser Asn Thr Thr Pro Glu Gly | | | | | | |
| | | 420 | | 425 | | 430 |
| Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr Glu Glu Ala Gly Glu Arg | | | | | | |
| | | 435 | | 440 | | 445 |
| Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn Leu Leu Lys Arg Leu Glu | | | | | | |
| | | 450 | | 455 | | 460 |
| Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg Glu Val Glu Arg Pro Leu | | | | | | |
| 465 | | 470 | | 475 | | 480 |
| Ser Ala Val Leu Ala His Met Glu Ala Thr Gly Val Arg Leu Asp Val | | | | | | |
| | | 485 | | 490 | | 495 |
| Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val Ala Glu Glu Ile Ala Arg | | | | | | |
| | | 500 | | 505 | | 510 |
| Leu Glu Ala Glu Val Phe Arg Leu Ala Gly His Pro Phe Asn Leu Asn | | | | | | |
| | | 515 | | 520 | | 525 |
| Ser Arg Asp Gln Leu Glu Arg Val Leu Phe Asp Glu Leu Gly Leu Pro | | | | | | |

| | | | | |
|---|-----|-----|-----|-----|
| 530 | | 535 | | 540 |
| Ala Ile Lys Lys Thr Gln Lys Thr Gly Lys Arg Ser Thr Ser Ala Ala | | | | |
| 545 | | 550 | | 555 |
| Val Leu Glu Ala Leu Arg Glu Ala His Pro Ile Val Glu Lys Ile Leu | | | | |
| | 565 | | 570 | 575 |
| Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser Thr Tyr Ile Asp Pro Leu | | | | |
| | 580 | | 585 | 590 |
| Pro Asp Leu Ile His Pro Arg Thr Gly Arg Leu His Thr Arg Phe Asn | | | | |
| | 595 | | 600 | 605 |
| Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser Ser Ser Asp Pro Asn Leu | | | | |
| | 610 | | 615 | 620 |
| Gln Asn Ile Pro Val Arg Thr Pro Leu Gly Gln Arg Ile Arg Arg Ala | | | | |
| | 625 | | 630 | 635 |
| Phe Ile Ala Glu Glu Gly Trp Leu Leu Val Ala Leu Asp Tyr Ser Gln | | | | |
| | 645 | | 650 | 655 |
| Ile Glu Leu Arg Val Leu Ala His Leu Ser Gly Asp Glu Asn Leu Ile | | | | |
| | 660 | | 665 | 670 |
| Arg Val Phe Gln Glu Gly Arg Asp Ile His Thr Glu Thr Ala Ser Trp | | | | |
| | 675 | | 680 | 685 |
| Met Phe Gly Val Pro Arg Glu Ala Val Asp Pro Leu Met Arg Arg Ala | | | | |
| | 690 | | 695 | 700 |
| Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr Gly Met Ser Ala His Arg | | | | |
| | 705 | | 710 | 715 |
| Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu Glu Ala Gln Ala Phe Ile | | | | |
| | 725 | | 730 | 735 |
| Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val Arg Ala Trp Ile Glu Lys | | | | |
| | 740 | | 745 | 750 |
| Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr Val Glu Thr Leu Phe Gly | | | | |
| | 755 | | 760 | 765 |
| Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala Arg Val Lys Ser Val Arg | | | | |
| | 770 | | 775 | 780 |
| Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro Val Gln Gly Thr Ala | | | | |
| | 785 | | 790 | 795 |
| Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu Phe Pro Arg Leu Glu | | | | |
| | 805 | | 810 | 815 |
| Glu Met Gly Ala Arg Met Leu Leu Gln Val Ala Asn Glu Leu Val Leu | | | | |
| | 820 | | 825 | 830 |
| Glu Ala Pro Lys Glu Arg Ala Glu Ala Val Ala Arg Leu Ala Lys Glu | | | | |

835 840 845
Val Met Glu Gly Val Tyr Pro Leu Ala Val Pro Leu Glu Val Glu Val
850 855 860
Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys Glu His His His His His
865 870 875 880

His

<210> 567
<211> 2643
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 567
atgggtgcgg atattggtga cctctttgag aggggaagagg tcgagcttga gtactttctca 60
ggaaagaaaa ttgccgttga tgctttcaac acgctatacc agttcatctc gataataagg 120
cagcctgacg gtacgccgtt aaaggactca cagggcagaa tcacctctca cctttccgga 180
atcctataca gagtctocaa catggctgag gtgggaatca ggccggtgtt tgtattcgac 240
ggagagccac cggagttcaa gaaggctgaa attgaggaga ggaaaaagag aagggtctgag 300
gcagaggaga tgtggattgc ggctttgcag gcaggagata aggacgcgaa aaagtatgct 360
caggctgcag ggaggggttga cgagtacatt gttgactccg caaagacgct ttttaagttac 420
atggggattc cctttgtcga tgccccgtct gaaggagagg cgcaggctgc ttacatggca 480
gcaaaaggcg atgtggagta cacaggaagc caggattacg attctctgct cttcgggaagc 540
ccgagactcg ccagaaatct cgcaataacg ggaaaaagga agcttcccgg caaaaatgtc 600
tatgtggatg taaagccgga gataataatt ctggaaagca acctcaaaag gctggggttg 660
acgagggagc agctcatcga catagcgatt ctggtcggga cggactacaa tgaggggtgtg 720
aagggtgtcg gcgtaagaa ggctttgaac tacatcaaga cctacggaga tattttcagg 780
gcactcaagg ctctgaaagt aaatattgac cacgtagagg agataaggaa tttcttcctg 840
aatcctcctg tgactgacga ctacagaata gagttcaggg agcctgactt tgagaaggcc 900
atcgagttcc tgtgagagga gcacgacttc agcagggaga gggtcgagaa ggccttgag 960
aagctcaaag ctctgaagtc aaccaggcc acgcttgaga ggtgggttct ggaggaggcc 1020
ccctggcccc cgccggaagg ggccttcgtg ggcttcgtcc tctcccgccc cgagcccatg 1080
tgggcggagc ttaaagccct ggccgctgc aggggcggcc gcgtccaccg ggcccccgag 1140

| | | | | | | |
|-------------|-------------|------------|------------|------------|------------|------|
| ccttataaaag | ccctcagggga | cctgaaggag | gcgcgggggc | ttctcgccaa | agacctgagc | 1200 |
| gttctggtccc | tgagggaagg | ccttggcctc | ccgcccggcg | acgaccccat | gtcctctgcc | 1260 |
| tacctcctgg | acccttcgaa | caccaccccc | gagggggtgg | cccggcgcta | cgcgggggag | 1320 |
| tggacggagg | aggcggggga | gcggggccgc | ctttccgaga | ggctcttcgc | caacctgtgg | 1380 |
| gggaggcttg | agggggagga | gaggtcctt | tggctttacc | gggaggtgga | gaggccctt | 1440 |
| tccgtgtcc | tggcccatat | ggaggccacg | ggggtgcgc | tggacgtggc | ctatctcagg | 1500 |
| gccttgtccc | tggaggtggc | cgaggagatc | gccgcctcg | aggccgaggt | cttcgcctg | 1560 |
| gccggccacc | ccttcaacct | caactcccgg | gaccagctgg | aaagggtcct | ctttgacgag | 1620 |
| ctagggcttc | ccgccatcgg | caagacggag | aagaccggca | agcgctccac | cagcgccgcc | 1680 |
| gtcctggagg | ccctccgcga | ggcccacccc | atcgtggaga | agatcctgca | gtaccgggag | 1740 |
| ctcaccaagc | tgaagagcac | ctacattgac | cccttgccgg | acctcatcca | ccccaggacg | 1800 |
| ggccgcctcc | acaccgcctt | caaccagacg | gccacggcca | cgggcaggct | aagtagctcc | 1860 |
| gatcccaacc | tccagaacat | ccccgtccgc | acccgccttg | ggcagaggat | ccgccggggc | 1920 |
| ttcatcgccg | aggaggggtg | gctattggtg | gccctggact | atagccagat | agagctcagg | 1980 |
| gtgctggccc | acctctccgg | cgacgagaac | ctgatccggg | tcttccagga | ggggcgggac | 2040 |
| atccacacgg | agaccgccag | ctggatgttc | ggcgtccccc | gggaggccgt | ggacccctg | 2100 |
| atgcgccggg | cggccaagac | catcaacttc | ggggtcctct | acggcatgtc | ggcccaccgc | 2160 |
| ctctcccagg | agctagccat | cccttacgag | gaggcccagg | ccttcattga | gcgctacttt | 2220 |
| cagagcttcc | ccaaggtgcg | ggcctggatt | gagaagacct | tggaggaggg | caggaggcgg | 2280 |
| gggtacgtgg | agaccctctt | cgcccgccgc | cgctacgtgc | cagacctaga | ggcccgggtg | 2340 |
| aagagcgtgc | gggaggcggc | cgagcgcacg | gccttcaaca | tgcccgtcca | gggcaccgcc | 2400 |
| gccgacctca | tgaagctggc | tatggtgaag | ctcttcccca | ggctggagga | aatggggggc | 2460 |
| aggatgctcc | ttcagggtcca | caacgagctg | gtcctcgagg | ccccaaaaga | gagggcgagg | 2520 |
| gccgtggccc | ggctggccaa | ggaggtcatg | gagggggtgt | atcccctggc | cgtgcccctg | 2580 |
| gaggtggagg | tggggatagg | ggaggactgg | ctctccgcca | aggagcacca | ccaccaccac | 2640 |
| cac | | | | | | 2643 |

<210> 568
 <211> 881
 <212> PRT

<213> Artificial

<220>

<223> Synthetic

<400> 568

Met Gly Ala Asp Ile Gly Asp Leu Phe Glu Arg Glu Glu Val Glu Leu
1 5 10 15
Glu Tyr Phe Ser Gly Lys Lys Ile Ala Val Asp Ala Phe Asn Thr Leu
20 25 30
Tyr Gln Phe Ile Ser Ile Ile Arg Gln Pro Asp Gly Thr Pro Leu Lys
35 40 45
Asp Ser Gln Gly Arg Ile Thr Ser His Leu Ser Gly Ile Leu Tyr Arg
50 55 60
Val Ser Asn Met Val Glu Val Gly Ile Arg Pro Val Phe Val Phe Asp
65 70 75 80
Gly Glu Pro Pro Glu Phe Lys Lys Ala Glu Ile Glu Glu Arg Lys Lys
85 90 95
Arg Arg Ala Glu Ala Glu Glu Met Trp Ile Ala Ala Leu Gln Ala Gly
100 105 110
Asp Lys Asp Ala Lys Lys Tyr Ala Gln Ala Ala Gly Arg Val Asp Glu
115 120 125
Tyr Ile Val Asp Ser Ala Lys Thr Leu Leu Ser Tyr Met Gly Ile Pro
130 135 140
Phe Val Asp Ala Pro Ser Glu Gly Glu Ala Gln Ala Ala Tyr Met Ala
145 150 155 160
Ala Lys Gly Asp Val Glu Tyr Thr Gly Ser Gln Asp Tyr Asp Ser Leu
165 170 175
Leu Phe Gly Ser Pro Arg Leu Ala Arg Asn Leu Ala Ile Thr Gly Lys
180 185 190
Arg Lys Leu Pro Gly Lys Asn Val Tyr Val Asp Val Lys Pro Glu Ile
195 200 205
Ile Ile Leu Glu Ser Asn Leu Lys Arg Leu Gly Leu Thr Arg Glu Gln
210 215 220
Leu Ile Asp Ile Ala Ile Leu Val Gly Thr Asp Tyr Asn Glu Gly Val
225 230 235 240
Lys Gly Val Gly Val Lys Lys Ala Leu Asn Tyr Ile Lys Thr Tyr Gly
245 250 255
Asp Ile Phe Arg Ala Leu Lys Ala Leu Lys Val Asn Ile Asp His Val
260 265 270

Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser Thr Tyr Ile Asp Pro Leu
580 585 590

Pro Asp Leu Ile His Pro Arg Thr Gly Arg Leu His Thr Arg Phe Asn
595 600 605

Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser Ser Ser Asp Pro Asn Leu
610 615 620

Gln Asn Ile Pro Val Arg Thr Pro Leu Gly Gln Arg Ile Arg Arg Ala
625 630 635 640

Phe Ile Ala Glu Glu Gly Trp Leu Leu Val Ala Leu Asp Tyr Ser Gln
645 650 655

Ile Glu Leu Arg Val Leu Ala His Leu Ser Gly Asp Glu Asn Leu Ile
660 665 670

Arg Val Phe Gln Glu Gly Arg Asp Ile His Thr Glu Thr Ala Ser Trp
675 680 685

Met Phe Gly Val Pro Arg Glu Ala Val Asp Pro Leu Met Arg Arg Ala
690 695 700

Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr Gly Met Ser Ala His Arg
705 710 715 720

Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu Glu Ala Gln Ala Phe Ile
725 730 735

Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val Arg Ala Trp Ile Glu Lys
740 745 750

Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr Val Glu Thr Leu Phe Gly
755 760 765

Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala Arg Val Lys Ser Val Arg
770 775 780

Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro Val Gln Gly Thr Ala
785 790 795 800

Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu Phe Pro Arg Leu Glu
805 810 815

Glu Met Gly Ala Arg Met Leu Leu Gln Val His Asn Glu Leu Val Leu
820 825 830

Glu Ala Pro Lys Glu Arg Ala Glu Ala Val Ala Arg Leu Ala Lys Glu
835 840 845

Val Met Glu Gly Val Tyr Pro Leu Ala Val Pro Leu Glu Val Glu Val
850 855 860

Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys Glu His His His His His
865 870 875 880

His

<210> 569
<211> 24
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 569
gagcggataa caatttcaca cagg 24

<210> 570
<211> 30
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 570
tgccccggtgc acgcggccgc ccctgcaggc 30

<210> 571
<211> 2445
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 571
atgaattccc tgccccctctt tgagcccaag ggccgggtgc ttctggtgga cggccaccac 60
ctggcctacc gtaccttttt tgccctgaag ggcctcacca ccagccgcgg ggagccggtc 120
caggcgggtgt acgggtttgc caagagcctt ttgaaggcgc taagggaaga cggggatgtg 180
gtgatcgtgg tgtttgacgc caaggccccc tccttccgcc accagaccta cgaggcctac 240
aaggcggggc gggctccac ccccgaggac tttcccggc agcttgccct tatcaaggag 300
atggtggacc ttttgggcct ggagcgcctc gaggtgccgg gctttgaagc ggatgacgtc 360
ctggctaccc tggccaagaa ggcggaaaag gagggctacg aggtccgcat cctcaccgcc 420
gacaaagacc tttaccagct cctttccgac cgcattccacg tcctccaccc cgaggggtac 480
ctcatcacc cggcctggct ttgggaaaag tacggcctga ggcccgacca gtgggcccgc 540
taccgggccc tgaccgggga cgagtccgac aaccttcccg gggtaaggg catcggggag 600

aaggaggcca tggagaaggc ctatcccctc gccgtgcccc tggaggtgga ggtggggatg 2400
 ggggaggact ggctttccgc caagggtcac caccaccacc accac 2445

<210> 572
 <211> 815
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 572

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Asn | Ser | Leu | Pro | Leu | Phe | Glu | Pro | Lys | Gly | Arg | Val | Leu | Leu | Val | 1 | 5 | 10 | 15 |
| Asp | Gly | His | His | Leu | Ala | Tyr | Arg | Thr | Phe | Phe | Ala | Leu | Lys | Gly | Leu | 20 | 25 | 30 | |
| Thr | Thr | Ser | Arg | Gly | Glu | Pro | Val | Gln | Ala | Val | Tyr | Gly | Phe | Ala | Lys | 35 | 40 | 45 | |
| Ser | Leu | Leu | Lys | Ala | Leu | Arg | Glu | Asp | Gly | Asp | Val | Val | Ile | Val | Val | 50 | 55 | 60 | |
| Phe | Asp | Ala | Lys | Ala | Pro | Ser | Phe | Arg | His | Gln | Thr | Tyr | Glu | Ala | Tyr | 65 | 70 | 75 | 80 |
| Lys | Ala | Gly | Arg | Ala | Pro | Thr | Pro | Glu | Asp | Phe | Pro | Arg | Gln | Leu | Ala | 85 | 90 | 95 | |
| Leu | Ile | Lys | Glu | Met | Val | Asp | Leu | Leu | Gly | Leu | Glu | Arg | Leu | Glu | Val | 100 | 105 | 110 | |
| Pro | Gly | Phe | Glu | Ala | Asp | Asp | Val | Leu | Ala | Thr | Leu | Ala | Lys | Lys | Ala | 115 | 120 | 125 | |
| Glu | Lys | Glu | Gly | Tyr | Glu | Val | Arg | Ile | Leu | Thr | Ala | Asp | Lys | Asp | Leu | 130 | 135 | 140 | |
| Tyr | Gln | Leu | Leu | Ser | Asp | Arg | Ile | His | Val | Leu | His | Pro | Glu | Gly | Tyr | 145 | 150 | 155 | 160 |
| Leu | Ile | Thr | Pro | Ala | Trp | Leu | Trp | Glu | Lys | Tyr | Gly | Leu | Arg | Pro | Asp | 165 | 170 | 175 | |
| Gln | Trp | Ala | Asp | Tyr | Arg | Ala | Leu | Thr | Gly | Asp | Glu | Ser | Asp | Asn | Leu | 180 | 185 | 190 | |
| Pro | Gly | Val | Lys | Gly | Ile | Gly | Glu | Lys | Thr | Ala | Arg | Lys | Leu | Leu | Glu | 195 | 200 | 205 | |
| Glu | Trp | Gly | Ser | Leu | Glu | Ala | Leu | Leu | Lys | Asn | Leu | Asp | Arg | Leu | Lys | 210 | 215 | 220 | |

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Pro | Ala | Ile | Arg | Glu | Lys | Ile | Leu | Ala | His | Met | Asp | Asp | Leu | Lys | Leu | 225 | 230 | 235 | 240 |
| Ser | Trp | Asp | Leu | Ala | Lys | Val | Arg | Thr | Asp | Leu | Pro | Leu | Glu | Val | Asp | 245 | 250 | 255 | |
| Phe | Ala | Lys | Arg | Arg | Glu | Pro | Asp | Arg | Glu | Gly | Glu | Lys | Pro | Arg | Glu | 260 | 265 | 270 | |
| Glu | Ala | Pro | Trp | Pro | Pro | Pro | Glu | Gly | Ala | Phe | Val | Gly | Phe | Leu | Leu | 275 | 280 | 285 | |
| Ser | Arg | Pro | Glu | Pro | Met | Trp | Ala | Glu | Leu | Lys | Ala | Leu | Ala | Ala | Cys | 290 | 295 | 300 | |
| Arg | Gly | Gly | Arg | Val | His | Arg | Ala | Ala | Asp | Pro | Leu | Ala | Gly | Leu | Lys | 305 | 310 | 315 | 320 |
| Asp | Leu | Lys | Glu | Val | Arg | Gly | Leu | Leu | Ala | Lys | Asp | Leu | Ala | Val | Leu | 325 | 330 | 335 | |
| Ala | Ser | Arg | Glu | Gly | Leu | Asp | Leu | Val | Pro | Gly | Asp | Asp | Pro | Met | Leu | 340 | 345 | 350 | |
| Leu | Ala | Tyr | Leu | Leu | Asp | Pro | Ser | Asn | Thr | Thr | Pro | Glu | Gly | Val | Ala | 355 | 360 | 365 | |
| Arg | Arg | Tyr | Gly | Gly | Glu | Trp | Thr | Glu | Asp | Ala | Ala | His | Arg | Ala | Leu | 370 | 375 | 380 | |
| Leu | Ser | Glu | Arg | Leu | His | Arg | Asn | Leu | Leu | Lys | Arg | Leu | Glu | Gly | Glu | 385 | 390 | 395 | 400 |
| Glu | Lys | Leu | Leu | Trp | Leu | Tyr | His | Glu | Val | Glu | Lys | Pro | Leu | Ser | Arg | 405 | 410 | 415 | |
| Val | Leu | Ala | His | Met | Glu | Ala | Thr | Gly | Val | Arg | Arg | Asp | Val | Ala | Tyr | 420 | 425 | 430 | |
| Leu | Gln | Ala | Leu | Ser | Leu | Glu | Leu | Ala | Glu | Glu | Ile | Arg | Arg | Leu | Glu | 435 | 440 | 445 | |
| Glu | Glu | Val | Phe | Arg | Leu | Ala | Gly | His | Pro | Phe | Asn | Leu | Asn | Ser | Arg | 450 | 455 | 460 | |
| Asp | Gln | Leu | Glu | Arg | Val | Leu | Phe | Asp | Glu | Leu | Arg | Leu | Pro | Ala | Leu | 465 | 470 | 475 | 480 |
| Lys | Lys | Thr | Lys | Lys | Thr | Gly | Lys | Arg | Ser | Thr | Ser | Ala | Ala | Val | Leu | 485 | 490 | 495 | |
| Glu | Ala | Leu | Arg | Glu | Ala | His | Pro | Ile | Val | Glu | Lys | Ile | Leu | Gln | His | 500 | 505 | 510 | |
| Arg | Glu | Leu | Thr | Lys | Leu | Lys | Asn | Thr | Tyr | Val | Asp | Pro | Leu | Pro | Ser | 515 | 520 | 525 | |

<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 573
atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
gaaccggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggccct caaggaggac 180
ggggacgcgg tgatcgtggt ctttgacgcc aaggccccct ccttcgccca cgaggcctac 240
gggggggtaca aggcggggccg ggccccacc ccggaggact tccccgccca gctcgccttg 300
gtcaagcggc tgggtggacct tctgggcctg gtccgcctcg agggcccggg gtacgaggcg 360
gacgacgtcc tgggcaccct ggccaagaag gccgaaaagg aggggtacga ggtgcgcata 420
ctcaccgccc accgcgacct ctaccaactc gtctccgacc gcatccacgt cctccacccc 480
tcatcacccc ggagtggctt tgggagaagt atgggcttaa gccttcccag 540
tgggtggact accgggcctt ggccggggac ccttcgcaca acatccccgg cgtgaagggc 600
atcggggaga agacggcggc caagctgata cgggagtggg gaagcctgga aaacctctc 660
aagaacctgg accggctgaa gcccgccatc cgggagaaga tcctggccca catggacgat 720
ctgaagctct cctgggacct ggccaagggt cgcaccgacc tggccctgga ggtggacttc 780
gccaaaaggc gggagcccga ccgggagagg cttagggcct ttctggagag gcttgagttt 840
ggcagcctcc tccacgagtt cggccttctg gaaagcccca aggcctgga ggaggcccc 900
tggccccgc cggaaggggc cttcgtgggc ttctctctct cccgccccga gcccatgtgg 960
gcggagctta aagccctggc cgcctgcagg ggcggccgcg tgcaccgggc agcagacccc 1020
ttggcggggc taaaggacct caaggaggtc cggggcctcc tcgccaagga cctcgcgctc 1080
ttggcctcga gggaggggct agacctcgtg cccggggacg acccatgct cctcgcctac 1140
ctcctggacc cttcgaacac ccccccgag ggggtggcgc ggcgctacgg gggggagtgg 1200
acggaggacg ccgcccaccg ggccctctct tcggagaggc tccatcgga cctccttaag 1260
cgctcagagg gggaggagaa gtccttttgg ctctaccacg aggtggaaaa gccctctcc 1320
cgggtcctgg cccatatgga ggccaccggg gtacggcggg acgtggccta ccttcaggcc 1380
ctttccctgg agcttgcgga ggagatccgc cgcctcagag aggaggtctt ccgcttggcg 1440
ggccacccct tcaacctcaa ctcccgggac cagctggaaa ggggtgctctt tgacgagctt 1500
aggcttcccc ccttgaagaa gacgaagaag acaggcaagc gctccaccag cgccgcggtg 1560

| | | | | | | |
|---|-----|----|--|-----|--|-----|
| 65 | | 70 | | 75 | | 80 |
| Gly Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg | | | | | | |
| | 85 | | | 90 | | 95 |
| Gln Leu Ala Leu Val Lys Arg Leu Val Asp Leu Leu Gly Leu Val Arg | | | | | | |
| | 100 | | | 105 | | 110 |
| Leu Glu Ala Pro Gly Tyr Glu Ala Asp Asp Val Leu Gly Thr Leu Ala | | | | | | |
| | 115 | | | 120 | | 125 |
| Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp | | | | | | |
| | 130 | | | 135 | | 140 |
| Arg Asp Leu Tyr Gln Leu Val Ser Asp Arg Ile His Val Leu His Pro | | | | | | |
| | 145 | | | 150 | | 155 |
| Glu Gly Tyr Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly Leu | | | | | | |
| | 165 | | | 170 | | 175 |
| Lys Pro Ser Gln Trp Val Asp Tyr Arg Ala Leu Ala Gly Asp Pro Ser | | | | | | |
| | 180 | | | 185 | | 190 |
| Asp Asn Ile Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Ala Lys | | | | | | |
| | 195 | | | 200 | | 205 |
| Leu Ile Arg Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu Asp | | | | | | |
| | 210 | | | 215 | | 220 |
| Arg Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp | | | | | | |
| | 225 | | | 230 | | 235 |
| Leu Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu | | | | | | |
| | 245 | | | 250 | | 255 |
| Glu Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg | | | | | | |
| | 260 | | | 265 | | 270 |
| Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly | | | | | | |
| | 275 | | | 280 | | 285 |
| Leu Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro | | | | | | |
| | 290 | | | 295 | | 300 |
| Glu Gly Ala Phe Val Gly Phe Val Leu Ser Arg Pro Glu Pro Met Trp | | | | | | |
| | 305 | | | 310 | | 315 |
| Ala Glu Leu Lys Ala Leu Ala Ala Cys Arg Gly Gly Arg Val His Arg | | | | | | |
| | 325 | | | 330 | | 335 |
| Ala Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val Arg Gly | | | | | | |
| | 340 | | | 345 | | 350 |
| Leu Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly Leu Asp | | | | | | |
| | 355 | | | 360 | | 365 |
| Leu Val Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro | | | | | | |

| | | |
|---|-----|-----|
| 675 | 680 | 685 |
| Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys | | |
| 690 | 695 | 700 |
| Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys Arg Gly | | |
| 705 | 710 | 715 |
| Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn | | |
| | 725 | 730 |
| Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn | | |
| | 740 | 745 |
| Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val | | |
| | 755 | 760 |
| Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu Leu Gln | | |
| | 770 | 775 |
| Val Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala Glu Glu | | |
| | 785 | 790 |
| Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro Leu Ala | | |
| | 805 | 810 |
| Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala | | |
| | 820 | 825 |
| Lys Gly His His His His His His | | |
| | 835 | 840 |

<210> 575
 <211> 2445
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

| | |
|--|-----|
| <400> 575 | |
| atgaattccc tgccctctt tgagcccaag ggccgggtgc ttctggtgga cggccaccac | 60 |
| ctggcctacc gtaccttttt tgccctgaag ggccctacca ccagccgcgg ggagccggtc | 120 |
| caggcgggtgt acgggtttgc caagagcctt ttgaaggcgc taagggaaga cggggatgtg | 180 |
| gtgatcgtgg tgtttgacgc caaggccccc tccttcgcgc accagacctt cgaggcctac | 240 |
| aaggcggggc gggctccac ccccgaggac tttcccggc agcttgccct tatcaaggag | 300 |
| atggtggacc ttttgggctt taccgcctc gaggtgccgg gctttgaagc ggatgacgtc | 360 |
| ctggctaccc tggccaagaa ggcggaag gagggctacg aggtccgcat cctcaccgcc | 420 |
| gacaaagacc tttaccagct cttttccgac cgcattccacg tcctccaccc cgaggggtac | 480 |

| | | | | | | |
|------------|------------|-------------|------------|-------------|------------|------|
| ctcatcacc | cggcctggct | ttgggaaaag | tacggcctga | ggcccagaca | gtgggcccag | 540 |
| taccggggcc | tgaccgggga | cgagtccgac | aaccttcccc | gggtcaagg | catcggggag | 600 |
| aagacggcga | ggaagcttct | ggaggagtgg | gggagcctgg | aagccctcct | caagaacctg | 660 |
| gaccggctga | agcccgccat | ccgggagaag | atcctggccc | acatggacga | tctgaagctc | 720 |
| tcttgggacc | tggccaaggt | gcgcaccgac | ctgcccctgg | aggtggactt | cgccaaaagg | 780 |
| cgggagcccc | accgggagg | ggagaagccc | cgggaggagg | ccccctggcc | cccgcgccga | 840 |
| ggggccttcg | tgggcttcct | cctttcccgc | cccagccca | tgtgggcgga | gcttaaagcc | 900 |
| ctggccgcct | gcaggggcgg | ccgcgtgcac | cgggcagcag | acccttggc | ggggctaaag | 960 |
| gacctcaagg | aggtccgggg | cctcctcgcc | aaggacctcg | ccgtcttggc | ctcgaggagg | 1020 |
| gggctagacc | tcgtgcccgg | ggacgacccc | atgctcctcg | cctacctcct | ggacccttcg | 1080 |
| aacaccaccc | ccgagggggg | ggcgcggcgc | tacggggggg | agtggacgga | ggacgcgcgc | 1140 |
| caccggggcc | tcctctcgga | gaggctccat | cggaacctcc | ttaagcgct | cgagggggag | 1200 |
| gagaagctcc | tttggctcta | ccacgaggtg | gaaaagcccc | tctcccggt | cctggcccat | 1260 |
| atggaggcca | ccggggtacg | gcgggacgtg | gcctaccttc | aggccctttc | cctggagctt | 1320 |
| gcggaggaga | tccgccgcct | cgaggaggag | gtcttccgct | tggcgggcca | ccccttcaac | 1380 |
| ctcaactccc | gggaccagct | ggaaaggggtg | ctctttgacg | agcttaggct | tccgccttg | 1440 |
| aagaagacga | agaagacagg | caagcgctcc | accagcgccg | cgggtgctgga | ggccctacgg | 1500 |
| gaggcccacc | ccatcgctga | gaagatcctc | cagcaccggg | agctcaccaa | gctcaagaac | 1560 |
| acctacgtgg | acccctccc | aagcctcgtc | cacccgagga | cgggcgcct | ccacaccgc | 1620 |
| ttcaaccaga | cggccacggc | cacggggagg | cttagtagct | ccgaccccaa | cctgcagaac | 1680 |
| atccccgtcc | gcacccctt | gggccagagg | atccgccggg | ccttcgtggc | cgaggcgggt | 1740 |
| tgggcgttgg | tggccctgga | ctatagccag | atagagctcc | gcgtcctcgc | ccacctctcc | 1800 |
| ggggacgaaa | acctgatcag | ggtcttccag | gaggggaagg | acatccacac | ccagaccgca | 1860 |
| agctggatgt | tcggcgtccc | cccggaggcc | gtggaccccc | tgatgcgccg | ggcgcccaag | 1920 |
| acggtgaact | tcggcgtcct | ctacggcatg | tccgcccata | ggctctccca | ggagcttgcc | 1980 |
| atcccctacg | aggaggcgg | ggcctttata | gagcgctact | tccaaagctt | ccccaaagg | 2040 |
| cgggcctgga | tagaaaagac | cctggaggag | gggaggaagc | ggggctacgt | ggaaacctc | 2100 |
| ttcggaaaga | ggcgctacgt | gcccagacct | aacgcccggg | tgaagagcgt | cagggaggcc | 2160 |
| gcggagcgca | tggccttcaa | catgcccgtc | cagggcaccg | ccgccgacct | catgaagctc | 2220 |

gccatggtga agctcttccc ccgcctccgg gagatggggg cccgcatgct cctccaggtc 2280
 gccaacgagc tcctcctgga ggccccccaa gcgcggggccg aggaggtggc ggctttggcc 2340
 aaggaggcca tggagaaggc ctatcccctc gccgtgcccc tggaggtgga ggtggggatg 2400
 ggggaggact ggctttccgc caagggtcac caccaccacc accac 2445

<210> 576
 <211> 815
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 576

Met Asn Ser Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu Val
 1 5 10 15
 Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu Lys Gly Leu
 20 25 30
 Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala Lys
 35 40 45
 Ser Leu Leu Lys Ala Leu Arg Glu Asp Gly Asp Val Val Ile Val Val
 50 55 60
 Phe Asp Ala Lys Ala Pro Ser Phe Arg His Gln Thr Tyr Glu Ala Tyr
 65 70 75 80
 Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu Ala
 85 90 95
 Leu Ile Lys Glu Met Val Asp Leu Leu Gly Phe Thr Arg Leu Glu Val
 100 105 110
 Pro Gly Phe Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys Lys Ala
 115 120 125
 Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys Asp Leu
 130 135 140
 Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu Gly Tyr
 145 150 155 160
 Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg Pro Asp
 165 170 175
 Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp Asn Leu
 180 185 190
 Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu Leu Glu

| | 805 | 810 | 815 | |
|-------|---|-----|-----|------|
| <210> | 577 | | | |
| <211> | 2520 | | | |
| <212> | DNA | | | |
| <213> | Artificial | | | |
| <220> | | | | |
| <223> | Synthetic | | | |
| <400> | 577 | | | |
| | atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac | | | 60 |
| | ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc | | | 120 |
| | gaaccggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggccct caaggaggac | | | 180 |
| | ggggacgcgg tgatcgtggt ctttgacgcc aaggccccct ccttcgcga cgaggcctac | | | 240 |
| | gggggggtaca aggcggggccg ggccccacc ccggaggact tccccgcca gctcgccttg | | | 300 |
| | gtcaagcggc tgggtggacct tctgggcttt accgcctcg agggcccggg gtacgaggcg | | | 360 |
| | gacgacgtcc tgggcaccct ggccaagaag gccgaaaagg aggggtacga ggtgcgcata | | | 420 |
| | ctcaccgccg accgcgacct ctaccaactc gtctccgacc gcatccacgt cctccacccc | | | 480 |
| | gaggggtacc tcatcacccc ggagtggctt tgggagaagt atgggcttaa gccttcccag | | | 540 |
| | tgggtggact accgggcctt ggccggggac ccttcgcga acatccccgg cgtgaagggc | | | 600 |
| | atcggggaga agacggcggc caagctgata cgggagtggg gaagcctgga aaacctctc | | | 660 |
| | aagaacctgg accggctgaa gcccgccatc cgggagaaga tcctggccca catggacgat | | | 720 |
| | ctgaagctct cctgggacct ggccaagggt cgcaccgacc tgcccctgga ggtggacttc | | | 780 |
| | gccaaaaggc gggagcccga ccgggagagg cttagggcct ttctggagag gcttgagttt | | | 840 |
| | ggcagcctcc tccacgagtt cggccttctg gaaagcccca aggcctgga ggaggcccc | | | 900 |
| | tggccccgc cggaaggggc cttcgtgggc ttcgtcctct cccgccccga gcccatgtgg | | | 960 |
| | gcggagctta aagccctggc cgctgcagg ggcggccgcg tgcaccgggc agcagacccc | | | 1020 |
| | ttggcggggc taaaggacct caaggaggtc cggggcctcc tcgccaagga cctcgcgctc | | | 1080 |
| | ttggcctcga gggaggggct agacctctg cccggggacg acccatgct cctcgcctac | | | 1140 |
| | ctcctggacc cttcgaacac ccccccgag ggggtggcgc ggcgctacgg gggggagtgg | | | 1200 |
| | acggaggacg ccgcccaccg ggccctctc tcggagaggc tccatcgga cctccttaag | | | 1260 |
| | cgctcagagg gggaggagaa gctcctttgg ctctaccacg aggtggaaaa gccctctcc | | | 1320 |
| | cgggtcctgg cccatatgga ggccaccggg gtacggcggg acgtggccta ccttcaggcc | | | 1380 |

ctttccctgg agcttgcgga ggagatccgc cgcctcgagg aggaggtctt ccgcttggcg 1440
ggccaccct tcaacctcaa ctcccgggac cagctggaaa ggggtgctctt tgacgagctt 1500
aggcttcccg ccttgaagaa gacgaagaag acaggcaagc gctccaccag cgcgcgggtg 1560
ctggaggccc tacgggaggc ccaccccatc gtggagaaga tctccagca ccgggagctc 1620
accaagctca agaacaccta cgtggacccc ctccaagcc tcgtccaccc gaggacgggc 1680
cgcctccaca cccgcttcaa ccagacggcc acggccacgg ggaggcttag tagctccgac 1740
cccaacctgc agaacatccc cgtccgcacc cccttggggc agaggatccg ccgggccttc 1800
gtggccgagg cgggttgggc gttggtggcc ctggactata gccagataga gctccgcgtc 1860
ctcgcccacc tctccgggga cgaaaacctg atcagggtct tccaggaggg gaaggacatc 1920
cacaccaga ccgcaagctg gatgttcggc gtcccccgagg aggccgtgga cccctgatg 1980
cgcggggcgg ccaagacggt gaacttcggc gtccctctacg gcatgtccgc ccataggctc 2040
tcccaggagc ttgccatccc ctacgaggag gcggtggcct ttatagagcg ctacttccaa 2100
agcttcccca aggtgcgggc ctggatagaa aagaccctgg aggaggggag gaagcggggc 2160
tacgtggaag ccctcttcgg aagaaggcgc tacgtgcccg acctcaacgc ccgggtgaag 2220
agcgtcaggg aggcgcgga gcgcattggc ttcaacatgc ccgtccaggg caccgcggcc 2280
gacctcatga agctcgccat ggtgaagctc ttccccgcgc tccgggagat gggggcccg 2340
atgctcctcc aggtcgccaa cgagctctc ctggaggccc cccaagcgcg ggccgaggag 2400
gtggcggctt tggccaagga ggccatggag aaggcctatc ccctcgccgt gcccctggag 2460
gtggaggtgg ggatggggga ggactggctt tccgccaagg gtcaccacca ccaccaccac 2520

<210> 578
<211> 840
<212> PRT
<213> Artificial

<220>
<223> Synthetic

<400> 578

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Asn | Ser | Glu | Ala | Met | Leu | Pro | Leu | Phe | Glu | Pro | Lys | Gly | Arg | Val |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Leu | Leu | Val | Asp | Gly | His | His | Leu | Ala | Tyr | Arg | Thr | Phe | Phe | Ala | Leu |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Lys | Gly | Leu | Thr | Thr | Ser | Arg | Gly | Glu | Pro | Val | Gln | Ala | Val | Tyr | Gly |
| | | 35 | | | | | 40 | | | | | 45 | | | |

Leu Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly Leu Asp
 355 360 365
 Leu Val Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro
 370 375 380
 Ser Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp
 385 390 395 400
 Thr Glu Asp Ala Ala His Arg Ala Leu Leu Ser Glu Arg Leu His Arg
 405 410 415
 Asn Leu Leu Lys Arg Leu Glu Gly Glu Glu Lys Leu Leu Trp Leu Tyr
 420 425 430
 His Glu Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met Glu Ala
 435 440 445
 Thr Gly Val Arg Arg Asp Val Ala Tyr Leu Gln Ala Leu Ser Leu Glu
 450 455 460
 Leu Ala Glu Glu Ile Arg Arg Leu Glu Glu Glu Val Phe Arg Leu Ala
 465 470 475 480
 Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu
 485 490 495
 Phe Asp Glu Leu Arg Leu Pro Ala Leu Lys Lys Thr Lys Lys Thr Gly
 500 505 510
 Lys Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His
 515 520 525
 Pro Ile Val Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys Leu Lys
 530 535 540
 Asn Thr Tyr Val Asp Pro Leu Pro Ser Leu Val His Pro Arg Thr Gly
 545 550 555 560
 Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu
 565 570 575
 Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu
 580 585 590
 Gly Gln Arg Ile Arg Arg Ala Phe Val Ala Glu Ala Gly Trp Ala Leu
 595 600 605
 Val Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu
 610 615 620
 Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys Asp Ile
 625 630 635 640
 His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu Ala Val
 645 650 655

Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly Val Leu
660 665 670

Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr
675 680 685

Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys
690 695 700

Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys Arg Gly
705 710 715 720

Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn
725 730 735

Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn
740 745 750

Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val
755 760 765

Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu Leu Gln
770 775 780

Val Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala Glu Glu
785 790 795 800

Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro Leu Ala
805 810 815

Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala
820 825 830

Lys Gly His His His His His His
835 840

<210> 579

<211> 2445

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 579

atgaattccc tgccctctt tgagcccaag ggccgggtgc ttctggtgga cggccaccac 60

ctggcctacc gtaccttttt tgccctgaag ggcctacca ccagccgcgg ggagccggtc 120

caggcgggtgt acgggtttgc caagagcctt ttgaaggcgc taagggaaga cggggatgtg 180

gtgatcgtgg tctttgacgc cgaggccccc tccttcgcc accagacctt cgaggcctac 240

aaggcggggc gggctccac ccccgaggac tttcccggc agcttgccct tatcaaggag 300

atggtggacc ttttggcct ggagcgctc gaggtgccgg gctttgaagc ggatgacgtc 360

| | | | | | | |
|-------------|-------------|------------|-------------|------------|------------|------|
| ctggctaccc | tggccaagaa | ggcggaaaag | gagggctacg | aggtccgcat | cctcaccgcc | 420 |
| gacaaagacc | tttaccagct | cctttccgac | cgcattccacg | tcctccaccc | cgaggggtac | 480 |
| ctcatcacc | cggcctggct | ttgggaaaag | tacggcctga | ggcccgaaca | gtgggcccga | 540 |
| taccgggccc | tgaccgggga | cgagtccgac | aaccttcccc | gggtcaagg | catcggggag | 600 |
| aagacggcga | ggaagcttct | ggaggagtgg | gggagcctgg | aagccctcct | caagaacctg | 660 |
| gaccggctga | agcccgccat | ccgggagaag | atcctggccc | acatggacga | tctgaagctc | 720 |
| tcctgggacc | tggccaaggt | gcgcaccgac | ctgcccctgg | aggtggactt | cgccaaaagg | 780 |
| cgggagcccg | accgggaggg | ggagaagccc | cgggaggagg | ccccctggcc | cccggccgaa | 840 |
| ggggccttcg | tgggcttcct | cctttcccg | cccagaccca | tgtgggcgga | gcttaaagcc | 900 |
| ctggccgct | gcaggggagg | ccgcgtgcac | cgggcagcag | accccttggc | ggggctaaag | 960 |
| gacctcaagg | aggtccgggg | cctcctcgcc | aaggacctcg | ccgtcttggc | ctcgagggag | 1020 |
| gggctagacc | tcgtgcccgg | ggacgacccc | atgctcctcg | cctacctcct | ggacccttcg | 1080 |
| aacaccaccc | ccgagggggg | ggcgcggcgc | tacggggggg | agtggacgga | ggacgccgcc | 1140 |
| caccgggccc | tcctctcgga | gaggctccat | cggaaacctc | ttaagcgct | cgagggggag | 1200 |
| gagaagctcc | tttggtctta | ccacgaggtg | gaaaagcccc | tctccgggt | cctggcccac | 1260 |
| atggaggcca | ccgggggtacg | gcgggacgtg | gcctaccttc | aggccctttc | cctggagctt | 1320 |
| gcggaggaga | tccgccgcct | cgaggaggag | gtcttcgcgt | tggcgggcca | ccccttcaac | 1380 |
| ctcaactccc | gggaccagct | ggaaagggtg | ctctttgacg | agcttaggct | tccgccttg | 1440 |
| aagaagacga | agaagacagg | caagcgctcc | accagcgccg | cggtgctgga | ggccctacgg | 1500 |
| gaggcccacc | ccatcggtga | gaagatcctc | cagcaccggg | agctcaccaa | gctcaagaac | 1560 |
| acctacgtgg | acccctccc | aagcctcgtc | caccgagga | cgggccgcct | ccacaccgc | 1620 |
| ttcaaccaga | cggccacggc | cacggggagg | cttagtagct | ccgaccccaa | cctgcagaac | 1680 |
| atccccgtcc | gcacccctt | gggacagagg | atccgcggg | ccttcgtggc | cgaggcgggt | 1740 |
| tgggcgttgg | tggccctgga | ctatagccag | atagagctcc | gcgtcctcgc | ccacctctcc | 1800 |
| ggggacgaaa | acctgatcag | ggtcttccag | gaggggaagg | acatccacac | ccagaccgca | 1860 |
| agctggatgt | tcggcgctcc | cccggaggcc | gtggaccccc | tgatgcgccg | ggcggccaag | 1920 |
| acggtgaact | tcggcgctcc | ctacggcatg | tccgcccata | ggctctccca | ggagcttgcc | 1980 |
| atccccctacg | aggaggcggg | ggcctttata | gagcgctact | tccaaagctt | ccccaaagtg | 2040 |

Glu Lys Ala Tyr Pro Leu Ala Val Pro Leu Glu Val Glu Val Gly Met
 785 790 795 800

Gly Glu Asp Trp Leu Ser Ala Lys Gly His His His His His His
 805 810 815

<210> 581
 <211> 2445
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 581
 atgaattccc tgcccctctt tgagcccaag ggccgggtgc ttctggtgga cggccaccac 60
 ctggcctacc gtaccttttt tgccctgaag ggccctacca ccagccgcgg ggagccggtc 120
 caggcgggtgt acgggtttgc caagagcctt ttgaaggcgc taagggaaga cggggatgtg 180
 gtgatcgtgg tctttgacgc cgaggccccc tcttccgcc accagacctt cgaggcctac 240
 aaggcggggc gggctccac ccccgaggac tttcccggc agcttgcct tatcaaggag 300
 atggtggacc ttttgggctt tacccgctc gaggtgccgg gctttgaagc ggatgacgtc 360
 ctggctaccc tggccaagaa ggcggaag gagggctacg aggtccgcat cctcaccgcc 420
 gacaaagacc ttaccagct cctttccgac cgcattcacg tctccacccc cgaggggtac 480
 ctcatcacc cggcctggct ttgggaaaag tacggcctga ggcccgacca gtgggccgac 540
 taccgggccc tgaccgggga cgagtccgac aaccttccc gggcaaggg catcggggag 600
 aagacggcga ggaagcttct ggaggagtgg gggagcctgg aagccctcct caagaacctg 660
 gaccggctga agcccgccat ccgggagaag atcctggccc acatggacga tctgaagctc 720
 tctgggacc tggccaaggt gcgcaccgac ctgcccctgg aggtggactt cgccaaaagg 780
 cgggagccc accgggaggg ggagaagccc cgggaggagg cccctggcc cccgcccga 840
 ggggccttcg tgggttctc cctttcccgc cccgagccca tgtgggcgga gcttaaagcc 900
 ctggccgcct gcaggggcgg ccgcgtgcac cgggcagcag accccttggc ggggctaaag 960
 gacctcaagg aggtccgggg cctcctcgcc aaggacctcg ccgtcttggc ctcgagggag 1020
 gggctagacc tcgtgcccgg ggaagacccc atgctcctcg cctacctcct ggacccttcg 1080
 aacaccaccc ccgagggggg ggcgcggcgc tacggggggg agtggacgga ggacgccgcc 1140
 caccgggccc tctctcggga gaggtccat cggaacctcc ttaagcgcct cgagggggag 1200
 gagaagctcc tttggctcta ccacgaggtg gaaaagcccc tctccgggt cctggcccat 1260

0936436-03404

atggaggcca ccggggtacg gcgggacgtg gcctaccttc aggccctttc cctggagctt 1320
 gcggaggaga tccgccgcct cgaggaggag gtcttccgct tggcgggcca ccccttcaac 1380
 ctcaactccc gggaccagct ggaaaggggtg ctctttgacg agcttaggct tcccgccttg 1440
 aagaagacga agaagacagg caagcgctcc accagcgccg cgggtgctgga ggccctacgg 1500
 gagggccacc ccatcgtgga gaagatcctc cagcaccggg agctcaccaa gctcaagaac 1560
 acctacgtgg accccctccc aagcctcgtc cacccgagga cgggccgcct ccacaccgcg 1620
 ttcaaccaga cggccacggc cacggggagg cttagtagct ccgaccccaa cctgcagaac 1680
 atccccgtcc gcaccccctt gggccagagg atccgccggg ccttcgtggc cgaggcgggt 1740
 tgggcgttgg tggccctgga ctatagccag atagagctcc gcgtcctcgc ccacctctcc 1800
 ggggacgaaa acctgatcag ggtcttccag gaggggaagg acatccacac ccagaccgca 1860
 agctggatgt tcggcgctcc cccggaggcc gtggaccccc tgatgcgccg ggcgccaag 1920
 acggtgaact tcggcgctct ctacggcatg tccgccata ggctctccca ggagcttgcc 1980
 atcccctacg aggaggcggg ggcctttata gagcgctact tccaaagctt ccccaagggtg 2040
 cgggcctgga tagaaaagac cctggaggag gggaggaagc ggggctacgt ggaaaccctc 2100
 ttcggaagaa ggcgctacgt gcccgcctc aacgcccggg tgaagagcgt cagggaggcc 2160
 gcggagcgca tggccttcaa catgcccgtc cagggcaccg ccgccgacct catgaagctc 2220
 gccatggtga agctcttccc ccgcctccgg gagatggggg ccgcgatgct cctccaggtc 2280
 gccaacgagc tcctcctgga ggccccccaa gcgcggggcc aggaggtggc ggctttggcc 2340
 aaggaggcca tggagaaggc ctatcccctc gccgtgcccc tggaggtgga ggtggggatg 2400
 ggggaggact ggctttccgc caagggtcac caccaccacc accac 2445

<210> 582
 <211> 815
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 582

Met Asn Ser Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu Val
 1 5 10 15

Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu Lys Gly Leu
 20 25 30

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Thr | Thr | Ser | Arg | Gly | Glu | Pro | Val | Gln | Ala | Val | Tyr | Gly | Phe | Ala | Lys | 35 | 40 | 45 | |
| Ser | Leu | Leu | Lys | Ala | Leu | Arg | Glu | Asp | Gly | Asp | Val | Val | Ile | Val | Val | 50 | 55 | 60 | |
| Phe | Asp | Ala | Glu | Ala | Pro | Ser | Phe | Arg | His | Gln | Thr | Tyr | Glu | Ala | Tyr | 65 | 70 | 75 | 80 |
| Lys | Ala | Gly | Arg | Ala | Pro | Thr | Pro | Glu | Asp | Phe | Pro | Arg | Gln | Leu | Ala | 85 | 90 | 95 | |
| Leu | Ile | Lys | Glu | Met | Val | Asp | Leu | Leu | Gly | Phe | Thr | Arg | Leu | Glu | Val | 100 | 105 | 110 | |
| Pro | Gly | Phe | Glu | Ala | Asp | Asp | Val | Leu | Ala | Thr | Leu | Ala | Lys | Lys | Ala | 115 | 120 | 125 | |
| Glu | Lys | Glu | Gly | Tyr | Glu | Val | Arg | Ile | Leu | Thr | Ala | Asp | Lys | Asp | Leu | 130 | 135 | 140 | |
| Tyr | Gln | Leu | Leu | Ser | Asp | Arg | Ile | His | Val | Leu | His | Pro | Glu | Gly | Tyr | 145 | 150 | 155 | 160 |
| Leu | Ile | Thr | Pro | Ala | Trp | Leu | Trp | Glu | Lys | Tyr | Gly | Leu | Arg | Pro | Asp | 165 | 170 | 175 | |
| Gln | Trp | Ala | Asp | Tyr | Arg | Ala | Leu | Thr | Gly | Asp | Glu | Ser | Asp | Asn | Leu | 180 | 185 | 190 | |
| Pro | Gly | Val | Lys | Gly | Ile | Gly | Glu | Lys | Thr | Ala | Arg | Lys | Leu | Leu | Glu | 195 | 200 | 205 | |
| Glu | Trp | Gly | Ser | Leu | Glu | Ala | Leu | Leu | Lys | Asn | Leu | Asp | Arg | Leu | Lys | 210 | 215 | 220 | |
| Pro | Ala | Ile | Arg | Glu | Lys | Ile | Leu | Ala | His | Met | Asp | Asp | Leu | Lys | Leu | 225 | 230 | 235 | 240 |
| Ser | Trp | Asp | Leu | Ala | Lys | Val | Arg | Thr | Asp | Leu | Pro | Leu | Glu | Val | Asp | 245 | 250 | 255 | |
| Phe | Ala | Lys | Arg | Arg | Glu | Pro | Asp | Arg | Glu | Gly | Glu | Lys | Pro | Arg | Glu | 260 | 265 | 270 | |
| Glu | Ala | Pro | Trp | Pro | Pro | Pro | Glu | Gly | Ala | Phe | Val | Gly | Phe | Leu | Leu | 275 | 280 | 285 | |
| Ser | Arg | Pro | Glu | Pro | Met | Trp | Ala | Glu | Leu | Lys | Ala | Leu | Ala | Ala | Cys | 290 | 295 | 300 | |
| Arg | Gly | Gly | Arg | Val | His | Arg | Ala | Ala | Asp | Pro | Leu | Ala | Gly | Leu | Lys | 305 | 310 | 315 | 320 |
| Asp | Leu | Lys | Glu | Val | Arg | Gly | Leu | Leu | Ala | Lys | Asp | Leu | Ala | Val | Leu | 325 | 330 | 335 | |

Ala Ser Arg Glu Gly Leu Asp Leu Val Pro Gly Asp Asp Pro Met Leu
340 345 350

Leu Ala Tyr Leu Leu Asp Pro Ser Asn Thr Thr Pro Glu Gly Val Ala
355 360 365

Arg Arg Tyr Gly Gly Glu Trp Thr Glu Asp Ala Ala His Arg Ala Leu
370 375 380

Leu Ser Glu Arg Leu His Arg Asn Leu Leu Lys Arg Leu Glu Gly Glu
385 390 395 400

Glu Lys Leu Leu Trp Leu Tyr His Glu Val Glu Lys Pro Leu Ser Arg
405 410 415

Val Leu Ala His Met Glu Ala Thr Gly Val Arg Arg Asp Val Ala Tyr
420 425 430

Leu Gln Ala Leu Ser Leu Glu Leu Ala Glu Glu Ile Arg Arg Leu Glu
435 440 445

Glu Glu Val Phe Arg Leu Ala Gly His Pro Phe Asn Leu Asn Ser Arg
450 455 460

Asp Gln Leu Glu Arg Val Leu Phe Asp Glu Leu Arg Leu Pro Ala Leu
465 470 475 480

Lys Lys Thr Lys Lys Thr Gly Lys Arg Ser Thr Ser Ala Ala Val Leu
485 490 495

Glu Ala Leu Arg Glu Ala His Pro Ile Val Glu Lys Ile Leu Gln His
500 505 510

Arg Glu Leu Thr Lys Leu Lys Asn Thr Tyr Val Asp Pro Leu Pro Ser
515 520 525

Leu Val His Pro Arg Thr Gly Arg Leu His Thr Arg Phe Asn Gln Thr
530 535 540

Ala Thr Ala Thr Gly Arg Leu Ser Ser Ser Asp Pro Asn Leu Gln Asn
545 550 555 560

Ile Pro Val Arg Thr Pro Leu Gly Gln Arg Ile Arg Arg Ala Phe Val
565 570 575

Ala Glu Ala Gly Trp Ala Leu Val Ala Leu Asp Tyr Ser Gln Ile Glu
580 585 590

Leu Arg Val Leu Ala His Leu Ser Gly Asp Glu Asn Leu Ile Arg Val
595 600 605

Phe Gln Glu Gly Lys Asp Ile His Thr Gln Thr Ala Ser Trp Met Phe
610 615 620

Gly Val Pro Pro Glu Ala Val Asp Pro Leu Met Arg Arg Ala Ala Lys
625 630 635 640

Thr Val Asn Phe Gly Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser
645 650 655

Gln Glu Leu Ala Ile Pro Tyr Glu Glu Ala Val Ala Phe Ile Glu Arg
660 665 670

Tyr Phe Gln Ser Phe Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu
675 680 685

Glu Glu Gly Arg Lys Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg
690 695 700

Arg Tyr Val Pro Asp Leu Asn Ala Arg Val Lys Ser Val Arg Glu Ala
705 710 715 720

Ala Glu Arg Met Ala Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp
725 730 735

Leu Met Lys Leu Ala Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met
740 745 750

Gly Ala Arg Met Leu Leu Gln Val Ala Asn Glu Leu Leu Leu Glu Ala
755 760 765

Pro Gln Ala Arg Ala Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met
770 775 780

Glu Lys Ala Tyr Pro Leu Ala Val Pro Leu Glu Val Glu Val Gly Met
785 790 795 800

Gly Glu Asp Trp Leu Ser Ala Lys Gly His His His His His His
805 810 815

<210> 583
<211> 2520
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 583
atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
ggccaccacc tggcctaccg tacctttttt gccctgaagg gcctcaccac cagccggggg 120
gagccggtcc aggcggtgta cgggtttgcc aagagccttt tgaaggcgct aagagaagac 180
ggggacgcgg tgatcgtggt ctttgacgcc gagggccct ccttcgcga cgaggcctac 240
ggggggtaca aggcggggcg ggctcccacc cccgaggact ttccccggca gcttgccctt 300
atcaaggagc tgggtggacct cctggggttt acccgctcg aggtccccgg ctacgaggcg 360
gacgacgttc tcgccaccct ggccaagaag gcggaaaagg aggggtacga ggtgcgcatc 420
ctcaccgccg acaaagacct ttaccagctc ctttcgcgacc gcatccacgt cctccacccc 480

| | |
|--|------|
| gaggggtacc tcatcacccc ggcttggtt tgggaaaagt acggcctgag gcccgaccag | 540 |
| tgggccgact accgggccct gaccggggac gagtccgaca accttcccgg ggtcaagggc | 600 |
| atcggggaga agaccgccct caagctcctc aaggagtggg ggagcctgga agccctcctc | 660 |
| aagaacctgg accggctgaa gcccggcatc cgggagaaga tcctggccca catggacgat | 720 |
| ctgaagctct cctgggacct ggccaaggtg cgcaccgacc tgccctgga ggtggacttc | 780 |
| gccaaaaggc gggagcccga ccgggagggg cttaaggcct ttttgagag gctggagttc | 840 |
| ggcagcctcc tccacgagtt cggcctcctg ggaggggaga agccccggga ggaggcccc | 900 |
| tggccccgc cggaaggggc ctctgtgggc tttgtgcttt cccgcaagga gcccatgtgg | 960 |
| gccgatcttc tggccctggc cgctgcagg ggcggccgcg tgcaccgggc agcagacccc | 1020 |
| ttggcggggc taaaggacct caaggaggtc cggggcctcc tcgccaagga cctcgccgtc | 1080 |
| ttggcctcga gggaggggct agacctcgtg cccggggacg accccatgct cctcgccctac | 1140 |
| ctcctggacc ctctgaacac ccccccgag ggggtggcgc ggcgtacgg gggggagtgg | 1200 |
| acggaggacg ccgcccaccg ggccctcctc tcggagaggc tccatcgga cctccttaag | 1260 |
| cgctcgagg gggaggagaa gctcctttgg ctctaccacg aggtggaaaa gccctctcc | 1320 |
| cgggtcctgg cccatatgga ggccaccggg gtacggcggg acgtggccta ccttcaggcc | 1380 |
| ctttccctgg agcttgcgga ggagatcgc cgctcgagg aggaggtctt ccgcttggcg | 1440 |
| ggccaccct tcaacctcaa ctccgggac cagctggaaa ggggtgcttt tgacgagctt | 1500 |
| aggcttccc ccttgaagaa gacgaagaag acaggcaagc gctccaccag cgccgcggtg | 1560 |
| ctggaggccc tacgggaggc ccacccatc gtggagaaga tcctccagca ccgggagctc | 1620 |
| accaagctca agaacaccta cgtggacccc ctcccaagcc tcgtccacc gaggacgggc | 1680 |
| cgctccaca cccgcttcaa ccagacggc acggccacgg ggaggcttag tagctccgac | 1740 |
| cccaacctgc agaacatccc cgtccgcacc cccttgggcc agaggatccg ccgggccttc | 1800 |
| gtggccgagg cgggttgggc gttggtggcc ctggactata gccagataga gctccgcgtc | 1860 |
| ctcgcccacc tctccgggga cgaaaacctg atcaggtct tccaggagg gaaggacatc | 1920 |
| cacaccaga ccgcaagctg gatgttcggc gtccccccgg aggccgtgga cccctgatg | 1980 |
| cgccgggcgg ccaagacggt gaacttcggc gtctctacg gcatgtccgc ccataggctc | 2040 |
| tcccaggagc ttgccatccc ctacgaggag gcggtggcct ttatagagcg ctacttccaa | 2100 |
| agcttcccca aggtgcgggc ctggatagaa aagacctgg aggaggggag gaagcggggc | 2160 |

Arg Val Phe Gln Glu Gly Lys Asp Ile His Thr Gln Thr Ala Ser Trp
610 615 620

Met Phe Gly Val Pro Pro Glu Ala Val Asp Pro Leu Met Arg Arg Ala
625 630 635 640

Ala Lys Thr Val Asn Phe Gly Val Leu Tyr Gly Met Ser Ala His Arg
645 650 655

Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu Glu Ala Val Ala Phe Ile
660 665 670

Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val Arg Ala Trp Ile Glu Lys
675 680 685

Thr Leu Glu Glu Gly Arg Lys Arg Gly Tyr Val Glu Thr Leu Phe Gly
690 695 700

Arg Arg Arg Tyr Val Pro Asp Leu Asn Ala Arg Val Lys Ser Val Arg
705 710 715 720

Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro Val Gln Gly Thr Ala
725 730 735

Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu Phe Pro Arg Leu Arg
740 745 750

Glu Met Gly Ala Arg Met Leu Leu Gln Val Ala Asn Glu Leu Leu Leu
755 760 765

Glu Ala Pro Gln Ala Arg Ala Glu Glu Val Ala Ala Leu Ala Lys Glu
770 775 780

Ala Met Glu Lys Ala Tyr Pro Leu Ala Val Pro Leu Glu Val Glu Val
785 790 795 800

Gly Met Gly Glu Asp Trp Leu Ser Ala Lys Gly
805 810

<210> 587
<211> 2433
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 587
atggaattcc tgccctctt tgagcccaag ggccgggtgc ttctggtgga cgccaccac 60
ctggcctacc gcaccttcca cgccctgaag ggcctcacca ccagccgcgg ggagccggtc 120
caggcgggtgt acgggtttgc caagagcctt ttgaaggcgc taagagaaga cggggatgtg 180
gtgatcgtgg tctttgacgc caaggccccc tccttcgcgc acgaggccta cgggggggtac 240

cttgccatcc cctacgagga ggcggtggcc tttatagagc gctacttcca aagcttcccc 2040
 aaggtgcggg cctggataga aaagaccctg gaggagggga ggaagcgggg ctacgtggaa 2100
 accctcttcg gaagaaggcg ctacgtgccc gacctcaacg cccgggtgaa gagcgtcagg 2160
 gaggccgcgg agcgcacatggc cttcaacatg cccgtccagg gcaccgccgc cgacctcatg 2220
 aagctcgcca tggatgaagct cttccccgcg ctccgggaga tggggggccg catgctcctc 2280
 caggtcgcca acgagctcct cctggaggcc ccccaagcgc gggccgagga ggtggcggct 2340
 ttggccaagg aggccatgga gaaggcctat cccctcgccg tgcccctgga ggtggaggtg 2400
 gggatggggg aggactggct ttccgccaag ggt 2433

<210> 588
 <211> 811
 <212> PRT
 <213> Artificial

<220>
 <223> Synthetic

<400> 588

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Glu | Phe | Leu | Pro | Leu | Phe | Glu | Pro | Lys | Gly | Arg | Val | Leu | Leu | Val | 1 | 5 | 10 | 15 |
| Asp | Gly | His | His | Leu | Ala | Tyr | Arg | Thr | Phe | His | Ala | Leu | Lys | Gly | Leu | 20 | 25 | 30 | |
| Thr | Thr | Ser | Arg | Gly | Glu | Pro | Val | Gln | Ala | Val | Tyr | Gly | Phe | Ala | Lys | 35 | 40 | 45 | |
| Ser | Leu | Leu | Lys | Ala | Leu | Arg | Glu | Asp | Gly | Asp | Val | Val | Ile | Val | Val | 50 | 55 | 60 | |
| Phe | Asp | Ala | Lys | Ala | Pro | Ser | Phe | Arg | His | Glu | Ala | Tyr | Gly | Gly | Tyr | 65 | 70 | 75 | 80 |
| Lys | Ala | Gly | Arg | Ala | Pro | Thr | Pro | Glu | Asp | Phe | Pro | Arg | Gln | Leu | Ala | 85 | 90 | 95 | |
| Leu | Ile | Lys | Glu | Leu | Val | Asp | Leu | Leu | Gly | Leu | Ala | Arg | Leu | Glu | Val | 100 | 105 | 110 | |
| Pro | Gly | Phe | Glu | Ala | Asp | Asp | Val | Leu | Ala | Thr | Leu | Ala | Lys | Lys | Ala | 115 | 120 | 125 | |
| Glu | Lys | Glu | Gly | Tyr | Glu | Val | Arg | Ile | Leu | Thr | Gly | Asp | Arg | Asp | Leu | 130 | 135 | 140 | |
| Tyr | Gln | Leu | Val | Ser | Asp | Arg | Val | Ala | Arg | Pro | Glu | Gln | Trp | Val | Asp | 145 | 150 | 155 | 160 |

Tyr Arg Ala Leu Ala Gly Asp Pro Ser Asp Asn Ile Pro Gly Val Lys
 165 170 175
 Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu Leu Glu Glu Trp Gly Ser
 180 185 190
 Val Glu Ala Leu Leu Lys Asn Leu Asp Arg Leu Lys Pro Ala Ile Arg
 195 200 205
 Glu Lys Ile Leu Ala His Met Glu Asp Leu Lys Leu Ser Leu Glu Leu
 210 215 220
 Ser Arg Val Arg Thr Asp Leu Pro Leu Glu Val Asp Leu Ala Gln Gly
 225 230 235 240
 Arg Glu Pro Asp Arg Glu Gly Leu Lys Ala Phe Leu Glu Arg Leu Glu
 245 250 255
 Phe Gly Ser Leu Leu His Glu Phe Gly Leu Leu Glu Ser Pro Val Ala
 260 265 270
 Ala Glu Glu Ala Pro Trp Pro Pro Glu Gly Ala Phe Val Gly Tyr
 275 280 285
 Val Leu Ser Arg Pro Glu Pro Met Trp Ala Glu Leu Asn Ala Leu Ala
 290 295 300
 Ala Ala Trp Gly Gly Arg Val His Arg Ala Ala Asp Pro Leu Ala Gly
 305 310 315 320
 Leu Lys Asp Leu Lys Glu Val Arg Gly Leu Leu Ala Lys Asp Leu Ala
 325 330 335
 Val Leu Ala Ser Arg Glu Gly Leu Asp Leu Val Pro Gly Asp Asp Pro
 340 345 350
 Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser Asn Thr Thr Pro Glu Gly
 355 360 365
 Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr Glu Asp Ala Ala His Arg
 370 375 380
 Ala Leu Leu Ser Glu Arg Leu His Arg Asn Leu Leu Lys Arg Leu Glu
 385 390 395 400
 Gly Glu Glu Lys Leu Leu Trp Leu Tyr His Glu Val Glu Lys Pro Leu
 405 410 415
 Ser Arg Val Leu Ala His Met Glu Ala Thr Gly Val Arg Arg Asp Val
 420 425 430
 Ala Tyr Leu Gln Ala Leu Ser Leu Glu Leu Ala Glu Glu Ile Arg Arg
 435 440 445
 Leu Glu Glu Glu Val Phe Arg Leu Ala Gly His Pro Phe Asn Leu Asn
 450 455 460

Ser Arg Asp Gln Leu Glu Arg Val Leu Phe Asp Glu Leu Arg Leu Pro
465 470 475 480

Ala Leu Lys Lys Thr Lys Lys Thr Gly Lys Arg Ser Thr Ser Ala Ala
485 490 495

Val Leu Glu Ala Leu Arg Glu Ala His Pro Ile Val Glu Lys Ile Leu
500 505 510

Gln His Arg Glu Leu Thr Lys Leu Lys Asn Thr Tyr Val Asp Pro Leu
515 520 525

Pro Ser Leu Val His Pro Arg Thr Gly Arg Leu His Thr Arg Phe Asn
530 535 540

Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser Ser Ser Asp Pro Asn Leu
545 550 555 560

Gln Asn Ile Pro Val Arg Thr Pro Leu Gly Gln Arg Ile Arg Arg Ala
565 570 575

Phe Val Ala Glu Ala Gly Trp Ala Leu Val Ala Leu Asp Tyr Ser Gln
580 585 590

Ile Glu Leu Arg Val Leu Ala His Leu Ser Gly Asp Glu Asn Leu Ile
595 600 605

Arg Val Phe Gln Glu Gly Lys Asp Ile His Thr Gln Thr Ala Ser Trp
610 615 620

Met Phe Gly Val Pro Pro Glu Ala Val Asp Pro Leu Met Arg Arg Ala
625 630 635 640

Ala Lys Thr Val Asn Phe Gly Val Leu Tyr Gly Met Ser Ala His Arg
645 650 655

Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu Glu Ala Val Ala Phe Ile
660 665 670

Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val Arg Ala Trp Ile Glu Lys
675 680 685

Thr Leu Glu Glu Gly Arg Lys Arg Gly Tyr Val Glu Thr Leu Phe Gly
690 695 700

Arg Arg Arg Tyr Val Pro Asp Leu Asn Ala Arg Val Lys Ser Val Arg
705 710 715 720

Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro Val Gln Gly Thr Ala
725 730 735

Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu Phe Pro Arg Leu Arg
740 745 750

Glu Met Gly Ala Arg Met Leu Leu Gln Val Ala Asn Glu Leu Leu Leu
755 760 765

Glu Ala Pro Gln Ala Arg Ala Glu Glu Val Ala Ala Leu Ala Lys Glu
770 775 780

Ala Met Glu Lys Ala Tyr Pro Leu Ala Val Pro Leu Glu Val Glu Val
785 790 795 800

Gly Met Gly Glu Asp Trp Leu Ser Ala Lys Gly
805 810

<210> 589
<211> 2493
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 589
atgaattccc tgccctcttt tgagcccaag ggccgggtcc tcctggtgga cggccaccac 60
ctggcctacc gtaccttttt tgccctgaag ggcctcacca cctcccgggg ggagccggtg 120
cagatggtct acggcttcgc ccggagcctc ctcaaggccc tcaaggagga cggggacgcg 180
gtgatcgtgg tctttgacgc cgaggccccc tccttcgcc accagaccta cgaggcctac 240
aaggcgggga gggctccac ccccgaggac tttcccggc agcttgccct tatcaaggag 300
ctggtggacc tcctgggggt taccgcctc gaggtcccc gctacgaggc ggacgacgtt 360
ctcgccaccc tggccaagaa ggcgaaaag gagggctacg aggtgcgcat cctcaccgcg 420
gaccgggacc tttaccagct tctttccgac cgcattcacg tccttcaccc cgaggggtac 480
ctcatcacc cggcctggct ttgggaaaag tacggcttga ggcccaccca gtgggcccgc 540
taccgggccc tgaccgggga cgaatccgac aacctttccg ggttcaaggg catcggggag 600
aagacggcga ggaagcttct ggaggagtgg gggagcctgg aagccctcct caagaacctg 660
gaccggctga agcccgccat ccgggagaag atcctggccc acatggacga tctgaagctc 720
tccttgagc tctcccgggt gcgcaccgac ctccccctgg aggtggactt cgccaaaagg 780
cgggagcccc accgggagag gcttagggcc tttctggaga ggcttgagtt tggcagcctc 840
ctccacgagt tcggcccttt ggaaagcccc agggccgcg aggaagctcc ctggccgccc 900
cccgaggag ccttcgtggg gtacgttctt tcccgcccc agcccatgtg ggcggagctt 960
aacgccttgg ccgccgccag gggcggccgc gtgcaccggg cagcagaccc cttggcgggg 1020
ctaaaggacc tcaaggaggt ccggggcctc ctgcgaagg acctcgccgt cttggcctcg 1080
agggaggggc tagacctcgt gcccggggac gaccccatgc tcctcgcccta cctcctggac 1140

```

ccttcgaaca ccacccccga gggggtggcg cggcgctacg ggggggagtg gacggaggac 1200
gccgcccacc gggccctcct ctcgagagag ctccatcgga acctccttaa gcgcctcgag 1260
ggggaggaga agctcctttg gctctaccac gaggtggaaa agccctctc cggggtcctg 1320
gcccatatgg aggccaccgg ggtacggcgg gacgtggcct accttcaggc cctttccctg 1380
gagcttgagg aggagatccg ccgcctcgag gaggaggtct tccgcttggc gggccacccc 1440
ttcaacctca actcccggga ccagctggaa aggggtgctct ttgacgagct taggcttccc 1500
gccttgaaga agacgaagaa gacaggcaag cgctccacca gcgccgcggg gctggaggcc 1560
ctacgggagg cccaccccat cgtggagaag atcctccagc accgggagct caccaagctc 1620
aagaacacct acgtggaccc cctcccaagc ctcgctccacc cgaggacggg ccgcctccac 1680
acccgcttca accagacggc cacggccacg gggaggctta gtagctccga ccccaacctg 1740
cagaacatcc ccgtccgcac ccccttgggc cagaggatcc gccgggcctt cgtggccgag 1800
gcgggttggg cgttgggtggc cctggactat agccagatag agtccgcgt cctcgcccac 1860
ctctccgggg acgaaaacct gatcagggtc ttccaggagg ggaaggacat ccacaccag 1920
accgcaagct ggatgttcgg cgtccccccg gaggccgtgg accccctgat gcgccgggcg 1980
gccaagacgg tgaacttcgg cgtcctctac ggcattgtcc cccataggct ctcccaggag 2040
cttgccatcc cctacgagga ggcgggtggc tttatagagc gctacttcca aagcttcccc 2100
aagggtgcgg cctggataga aaagacctg gaggagggga ggaagcgggg ctacgtggaa 2160
acctctctcg gaagaaggcg ctacgtgccc gacctcaacg cccgggtgaa gagcgtcagg 2220
gaggccgcgg agcgcattgg cttcaacatg cccgtccagg gcaccgcgcg cgacctcatg 2280
aagctcgcca tgggtgaagc cttccccgcg ctccgggaga tgggggcccg catgctcctc 2340
caggctcgcca acgagctcct cctggaggcc cccaagcgc gggccgagga ggtggcggct 2400
ttggccaagg aggccatgga gaaggcctat cccctcgccg tgcccctgga ggtggagggtg 2460
gggatggggg aggactggct ttccgccaag ggt 2493

```

```

<210> 590
<211> 831
<212> PRT
<213> Artificial

```

```

<220>
<223> Synthetic

```

```

<400> 590

```


Phe Val Gly Tyr Val Leu Ser Arg Pro Glu Pro Met Trp Ala Glu Leu
 305 310 315 320
 Asn Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala Ala Asp
 325 330 335
 Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val Arg Gly Leu Leu Ala
 340 345 350
 Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly Leu Asp Leu Val Pro
 355 360 365
 Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser Asn Thr
 370 375 380
 Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr Glu Asp
 385 390 395 400
 Ala Ala His Arg Ala Leu Leu Ser Glu Arg Leu His Arg Asn Leu Leu
 405 410 415
 Lys Arg Leu Glu Gly Glu Glu Lys Leu Leu Trp Leu Tyr His Glu Val
 420 425 430
 Glu Lys Pro Leu Ser Arg Val Leu Ala His Met Glu Ala Thr Gly Val
 435 440 445
 Arg Arg Asp Val Ala Tyr Leu Gln Ala Leu Ser Leu Glu Leu Ala Glu
 450 455 460
 Glu Ile Arg Arg Leu Glu Glu Glu Val Phe Arg Leu Ala Gly His Pro
 465 470 475 480
 Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe Asp Glu
 485 490 495
 Leu Arg Leu Pro Ala Leu Lys Lys Thr Lys Lys Thr Gly Lys Arg Ser
 500 505 510
 Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro Ile Val
 515 520 525
 Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys Leu Lys Asn Thr Tyr
 530 535 540
 Val Asp Pro Leu Pro Ser Leu Val His Pro Arg Thr Gly Arg Leu His
 545 550 555 560
 Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser Ser Ser
 565 570 575
 Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly Gln Arg
 580 585 590
 Ile Arg Arg Ala Phe Val Ala Glu Ala Gly Trp Ala Leu Val Ala Leu
 595 600 605

<213> Artificial

<220>

<223> Synthetic

<400> 592
 tggaggtcaa aacatcgata agtcgaagaa aggaagggaa gaaat 45

<210> 593
 <211> 14
 <212> DNA
 <213> Artificial

<220>

<223> Synthetic

<400> 593
 tgttttgacc tcca 14

<210> 594
 <211> 20
 <212> DNA
 <213> Artificial

<220>

<223> Synthetic

<400> 594
 ggcgaccaca cccgtcctgt 20

<210> 595
 <211> 20
 <212> DNA
 <213> Artificial

<220>

<223> Synthetic

<400> 595
 ccacgatgcg tccggcgtag 20

<210> 596
 <211> 26
 <212> DNA
 <213> Artificial

<220>

<223> Synthetic

<400> 596
 acgggtcaat gtccatgccc caaaga 26

<210> 597
 <211> 41
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <220>
 <221> modified_base
 <222> (23)..(23)
 <223> /note= "The residue at this position indicates 2' o-methyl sugar"

<400> 597
 gtctgagatg aaagtgtcc cgcacccacc caaggcacag c 41

<210> 598
 <211> 26
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 598
 tcttcgcaca tttcatctca gacgga 26

<210> 599
 <211> 17
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<220>
 <221> modified_base
 <222> (1)..(17)
 <223> /note= "The residues at these positions are 2' o-methyl sugars"

<400> 599
 gctgtgcctt gggtggg 17

<210> 600
 <211> 21
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<220>
 <221> modified_base
 <222> (1)..(21)
 <223> /note= "The residues at these positions are 2' o-methyls"

<400> 600
 gctgtgcctt gggtaggtgc g 21

<210> 601
 <211> 28
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 601
 aacgaggcgc acccacccaa ggcacagc 28

<210> 602
 <211> 27
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<220>
 <221> modified_base
 <222> (23)..(27)
 <223> /note= "The residues at these positions are 2' o-methyls"

<400> 602
 gtctgagatg aaagtgcgcc tcgttaa 27

<210> 603
 <211> 28
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 603
 aacgaggcgc acccacccaa ggcacagc 28

<210> 604
 <211> 23
 <212> DNA
 <213> Artificial

```

<220>
<223> Synthetic

<400> 604
tcttcgcaca tttcatctca gac
23

<210> 605
<211> 21
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 605
gctgtgcctt gggtaggtgc g
21

<210> 606
<211> 26
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 606
tcttcgcaca tttcatctca gacgga
26

<210> 607
<211> 20
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 607
gctgtgcctt gggtaggtgc
20

<210> 608
<211> 20
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<220>
<221> modified_base
<222> (15)..(20)
<223> /note= "The residues at these positions are 2' o-methyl sugars"

```

<400> 608
gctgtgcctt ggggtgggtgc 20

<210> 609
<211> 20
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<220>
<221> modified_base
<222> (1)..(20)
<223> /note= "The residues at these positions are 2' o-methyl sugars"

<400> 609
gctgtgcctt ggggtgggtgc 20

<210> 610
<211> 26
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 610
acgggtcaat gtccatgccc caaaga 26

<210> 611
<211> 28
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 611
aacgaggcgc acccacccaa ggcacagc 28

<210> 612
<211> 27
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<220>
<221> modified_base
<222> (23)..(27)

<223> /note= "The residues at these positions are 2' o-methyls"

<400> 612
gtctgagatg aaagtgcgcc tcgttaa

27

<210> 613
<211> 17
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<220>
<221> modified_base
<222> (1)..(17)
<223> /note= "The residues at these positions are 2' o-methyls"

<400> 613
gctgtgcctt gggtggg

17

<210> 614
<211> 19
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<220>
<221> modified_base
<222> (1)..(19)
<223> /note= "The residues at these positions are 2' o-methyls"

<400> 614
gctgtgcctt gggtgggtg

19

<210> 615
<211> 21
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<220>
<221> modified_base
<222> (1)..(21)
<223> /note= "The residues at these positions are 2' o-methyls"

| | |
|--|----|
| <400> 615 gctgtgcctt gggtaggtgc g | 21 |
| | |
| <210> 616 <211> 22 <212> DNA <213> Artificial | |
| <220> <223> Synthetic | |
| <220> <221> modified_base <222> (1)..(22) <223> /note= "The residues at these positions are 2' o-methyls" | |
| | |
| <400> 616 gctgtgcctt gggtaggtgc gc | 22 |
| | |
| <210> 617 <211> 26 <212> DNA <213> Artificial | |
| <220> <223> Synthetic | |
| | |
| <400> 617 ccgtcaacat ttaccatggg tgcgga | 26 |
| | |
| <210> 618 <211> 31 <212> DNA <213> Artificial | |
| <220> <223> Synthetic | |
| | |
| <400> 618 ccgccacctc gtagtcgaca tccttttcgt g | 31 |
| | |
| <210> 619 <211> 29 <212> DNA <213> Artificial | |
| <220> <223> Synthetic | |
| | |
| <400> 619 ttgataaatt tgggggtggaa aggtttgga | 29 |

<400> 624
atgagtgagt ttggtgcg 18

<210> 625
<211> 27
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 625
ccaggaagca agtggtgcgc ctcgttt 27

<210> 626
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 626
aacgaggcgc accttggagg ca 22

<210> 627
<211> 24
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 627
aaggtttcct ttcagttgt gtta 24

<210> 628
<211> 26
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 628
gcaaagatgt ctgttacggt caactc 26

<210> 629
<211> 16
<212> DNA
<213> Artificial

<220>
 <223> Synthetic
 <400> 629
 tgcctccaag gtgcgc 16

<210> 630
 <211> 27
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic
 <400> 630
 aacgaggcgc accttcaaaa tgcctaa 27

<210> 631
 <211> 23
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic
 <400> 631
 tgtcactctc ctctttccaa tta 23

<210> 632
 <211> 29
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic
 <400> 632
 gaaaagagtt ccattatccg ctacatctg 29

<210> 633
 <211> 21
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic
 <400> 633
 ttaggcattt tgaaggtgcg c 21

<210> 634
 <211> 23

<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 634
aacgaggcgc accgttgtgt ccc 23

<210> 635
<211> 21
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 635
gggatgtaga agccattcag a 21

<210> 636
<211> 20
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 636
ttgtttgtgct gtgggggatg 20

<210> 637
<211> 17
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 637
gggacacaac ggtgcgc 17

<210> 638
<211> 23
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 638
ccgtcacgcc tccaccatat ccc 23

<210> 639
<211> 27
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 639
ccagcggttt ccattggcaa agatcaa

27

<210> 640
<211> 21
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 640
cggaagaatg ggtagacat g

21

<210> 641
<211> 17
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 641
gggatatggt ggagggcg

17

<210> 642
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 642
aacgaggcgc accgttccag gc

22

<210> 643
<211> 24
<212> DNA
<213> Artificial

<220>
<223> Synthetic

| | |
|--|----|
| <400> 643 catatccatg cagcaccacc atga | 24 |
| <210> 644 <211> 25 <212> DNA <213> Artificial | |
| <220> <223> Synthetic | |
| <400> 644 caaaatacag agtgaacaca gggcc | 25 |
| <210> 645 <211> 16 <212> DNA <213> Artificial | |
| <220> <223> Synthetic | |
| <400> 645 gcctggaacg gtgcgc | 16 |
| <210> 646 <211> 25 <212> DNA <213> Artificial | |
| <220> <223> Synthetic | |
| <400> 646 ccgtcacgcc tcatggataa tgccc | 25 |
| <210> 647 <211> 33 <212> DNA <213> Artificial | |
| <220> <223> Synthetic | |
| <400> 647 caggtgagaa aaggcattac agatagtgaa agc | 33 |
| <210> 648 <211> 22 <212> DNA <213> Artificial | |

<212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 653
 tttctcagca gggaggcg 18

 <210> 654
 <211> 20
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 654
 ccgtcacgcc tcgccccaca 20

 <210> 655
 <211> 28
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 655
 cagcacaggc tggtgacat cataaaac 28

 <210> 656
 <211> 25
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 656
 cttttccata ctttttatga cattc 25

 <210> 657
 <211> 14
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 657
 tgtggggcga ggcg 14

<210> 658
 <211> 23
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 658
 aacgaggcgc acagttgacc ttc

23

<210> 659
 <211> 19
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 659
 gtgatggcca gcacagggc

19

<210> 660
 <211> 20
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 660
 atacgttccc cacatttttc

20

<210> 661
 <211> 18
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 661
 tgaaggtcaa ctgtgcgc

18

<210> 662
 <211> 26
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

660 661 662 663 664 665 666 667

<400> 662
aacgaggcgc acgtcataaa tacccc 26

<210> 663
<211> 21
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 663
gccagcatag gctgttgaca c 21

<210> 664
<211> 28
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 664
agacttttct atacttttta taacattc 28

<210> 665
<211> 20
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 665
ggggtattta tgacgtgcgc 20

<210> 666
<211> 23
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 666
ccgtcacgcc tectgtctgt gat 23

<210> 667
<211> 22
<212> DNA
<213> Artificial

<220>
 <223> Synthetic

 <400> 667
 tcctgacagt gctcaatcag ga 22

 <210> 668
 <211> 22
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 668
 tcctgacaat gctcaatgag ga 22

 <210> 669
 <211> 16
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 669
 gtcccggatg tggccc 16

 <210> 670
 <211> 17
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 670
 atcacagaca ggaggcg 17

 <210> 671
 <211> 26
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 671
 aacgaggcgc acggactggt ttctgc 26

 <210> 672
 <211> 27

<212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 672
 cttgtcaaag tcctgatagt gtcctc 27

 <210> 673
 <211> 27
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 673
 cttgttgaag tcttgatagt gttcctc 27

 <210> 674
 <211> 20
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 674
 gcagaaaaca gtccgtgcgc 20

 <210> 675
 <211> 23
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 675
 ccgtcacgcc tcaactgcggt cat 23

 <210> 676
 <211> 31
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 676
 gtggataact gcatcagtggt atggcatttt c 31

| | | |
|-------|-------------------------------|----|
| <210> | 677 | |
| <211> | 25 | |
| <212> | DNA | |
| <213> | Artificial | |
| <220> | | |
| <223> | Synthetic | |
| <400> | 677 | 25 |
| | caagggttg tagcctgtgt gagcc | |
| <210> | 678 | |
| <211> | 17 | |
| <212> | DNA | |
| <213> | Artificial | |
| <220> | | |
| <223> | Synthetic | |
| <400> | 678 | 17 |
| | atgaccgcag tgaggcg | |
| <210> | 679 | |
| <211> | 24 | |
| <212> | DNA | |
| <213> | Artificial | |
| <220> | | |
| <223> | Synthetic | |
| <400> | 679 | 24 |
| | ccgtcacgcc tcagagccaa tcac | |
| <210> | 680 | |
| <211> | 27 | |
| <212> | DNA | |
| <213> | Artificial | |
| <220> | | |
| <223> | Synthetic | |
| <400> | 680 | 27 |
| | cgatcatcaa gggatggtgg cctgtgc | |
| <210> | 681 | |
| <211> | 31 | |
| <212> | DNA | |
| <213> | Artificial | |
| <220> | | |
| <223> | Synthetic | |

<220>
 <223> Synthetic

<400> 686
 cagaaattga agaggaggcg 20

<210> 687
 <211> 21
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 687
 aacgaggcgc accgggtccc a 21

<210> 688
 <211> 30
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 688
 tccccctgttt cttgaaaagt ccatgtgtga 30

<210> 689
 <211> 21
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 689
 aatccgtaga ggagcaccag g 21

<210> 690
 <211> 15
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 690
 tgggacccgg tgcgc 15

<210> 691
 <211> 21

| | |
|---------------------------------|----|
| <212> DNA | |
| <213> Artificial | |
| <220> | |
| <223> Synthetic | |
| <400> 691 | 21 |
| ccgtcacgcc tcctcggcag g | |
| <210> 692 | |
| <211> 29 | |
| <212> DNA | |
| <213> Artificial | |
| <220> | |
| <223> Synthetic | |
| <400> 692 | 29 |
| cacaatatcg taggtaggag gtgccttaa | |
| <210> 693 | |
| <211> 18 | |
| <212> DNA | |
| <213> Artificial | |
| <220> | |
| <223> Synthetic | |
| <400> 693 | 18 |
| gccccatcga tctcctcc | |
| <210> 694 | |
| <211> 15 | |
| <212> DNA | |
| <213> Artificial | |
| <220> | |
| <223> Synthetic | |
| <400> 694 | 15 |
| cctgccgagg aggcg | |
| <210> 695 | |
| <211> 23 | |
| <212> DNA | |
| <213> Artificial | |
| <220> | |
| <223> Synthetic | |
| <400> 695 | 23 |
| aacgaggcgc actaggcttt gct | |

<210> 696
 <211> 31
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 696
 ttcatgtagt cagggtcata gacaattaag a 31

 <210> 697
 <211> 23
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 697
 tccccagaac catcgaggaa agg 23

 <210> 698
 <211> 17
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 698
 agcaaagcct agtgcg 17

 <210> 699
 <211> 24
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

 <400> 699
 aacgaggcg acagaaggcc cctt 24

 <210> 700
 <211> 31
 <212> DNA
 <213> Artificial

 <220>
 <223> Synthetic

<400> 700
 ccttgaacag caccagaaat agactgagca c 31

<210> 701
 <211> 23
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 701
 ggaagaaccc agagacacca tcc 23

<210> 702
 <211> 18
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 702
 aaggggcctt ctgtgcgc 18

<210> 703
 <211> 25
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 703
 aacgaggcgc acgttgtgat acctt 25

<210> 704
 <211> 27
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 704
 gatgaaggcc ataaattaaa attgtgc 27

<210> 705
 <211> 16
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic
 <400> 705
 tgggtatgga acgtcc 16

<210> 706
 <211> 19
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic
 <400> 706
 aaggtatcac aacgtgcgc 19

<210> 707
 <211> 14
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic
 <400> 707
 cccccttttg gggg 14

<210> 708
 <211> 30
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic
 <400> 708
 ccctatcttt aaagttttta aaaagtttga 30

1106